THE DIFFERENTIATION OF RESPONSE NUMEROSITIES IN THE PIGEON

ARMANDO MACHADO[1] AND PAULO RODRIGUES[2]

UNIVERSITY OF MINHO, PORTUGAL,[1] AND
UNIVERSITY OF BEIRA INTERIOR, PORTUGAL[2]

Two experiments examined how pigeons differentiate response patterns along the dimension of number. In Experiment 1, 5 pigeons received food after pecking the left key at least N times and then switching to the right key (Mechner's Fixed Consecutive Number schedule). Parameter N varied across conditions from 4 to 32. Results showed that run length on the left key followed a normal distribution whose mean and standard deviation increased linearly with N; the coefficient of variation approached a constant value (the scalar property). In Experiment 2, 4 pigeons received food with probability p for pecking the left key exactly four times and then switching. If that did not happen, the pigeons still could receive food by returning to the left key and pecking it for a total of at least 16 times and then switching. Parameter p varied across conditions from 1.0 to .25. Results showed that when $p=1.0$ or $p=.5$, pigeons learned two response numerosities within the same condition. When $p=.25$, each pigeon adapted to the schedule differently. Two of them emitted first runs well described by a mixture of two normal distributions, one with mean close to 4 and the other with mean close to 16 pecks. A mathematical model for the differentiation of response numerosity in Fixed Consecutive Number schedules is proposed.

Key words: Fixed Consecutive Number schedule, numerosity differentiation, quantitative model, scalar property, key peck, pigeon

———————————

A growing literature supports the idea that number is a fundamental dimension of stimulus control (for reviews, see Emmerton, 2001; Gallistel, 1990; Rilling, 1993; Roberts, 1998). Thus rats and pigeons can learn to choose alternative A or B according to the number of *stimuli* presented to them. For example, a rat can learn to press a left lever following a sequence of two tones and press a right lever following a sequence of eight tones (Meck & Church, 1983; also Alsop & Honig, 1991; Keen & Machado, 1999; Machado & Keen, 2002). Animals also can learn to choose alternative A or B according to the number of *responses* emitted during a previous period. In a typical arrangement, a pigeon pecks repeatedly at a center key. After 35 or 50 pecks (decided randomly) that key is turned off, and two side keys are illuminated. To receive food the pigeon must choose the key on the left side if it pecked the center key 35 times and the key on the right side if it pecked the center key 50 times (e.g., Rilling, 1965, 1967; also Fetterman, 1993). The results from these studies show that

animals can learn to discriminate at least two numerosities of a single stimulus or response.

In the preceding studies, the experimenter, not the animal, controlled the sample numerosities. In one case the experimenter presented the stimulus for a predetermined number of times, and in the other case the experimenter prevented the animal from continuing to respond (by turning off the key light or removing the lever as the case may be) after a predetermined number of times. In both cases the (stimulus or response) numerosity was an independent variable whose values were under the experimenter's control. In contrast, in other studies the animal, not the experimenter, controls the numerosity. To illustrate, in Mechner's (1958) study, rats received food for pressing a left lever at least N times and then switching to a right lever; if they switched to the right lever before meeting the criterion of N responses on the left lever, then the trial was cancelled. The rats learned to switch to the right lever only after producing runs of approximately N responses on the left lever. In this Fixed Consecutive Number (FCN) schedule, the number of responses on the left lever (the run length) was a dependent variable whose values were under the animal's control.

Another difference between the two sets of studies is that the former involved relative

numerosity (e.g., two versus eight tones, or 35 versus 50 pecks) whereas the latter involved absolute numerosity (N or more presses on the left lever before switching). To distinguish them, we refer to the first set as numerosity *discrimination* and to the second set as numerosity *differentiation.*

Although procedurally different, discrimination and differentiation along the dimension of number may depend on common processes such as a neural counting mechanism (Gallistel, 1990; Meck & Church, 1983; Meck, Church, & Gibbon, 1985; Roberts, 1995; Roberts & Mitchell, 1994; but see Roberts, Roberts, & Kit, 2002). According to one version of such a mechanism, a pacemaker feeds pulses into an accumulator during a brief period of time (e.g., 200 ms) initiated by each occurrence of the stimulus or response whose numerosity the animal must discriminate or differentiate. At the moment of reinforcement, the pulses in the accumulator are transferred to a memory store. To choose between left and right alternatives in discrimination procedures, or to decide when to switch to the right lever in differentiation procedures, the animal extracts a sample from the memory store at trial onset and then compares the value of that sample with the value in the accumulator; the decision rule is based on the similarity between the two numbers (for further details, see, e.g., Meck & Church, 1983; Meck, Church, & Gibbon, 1985).

The hypothesis of a common neural mechanism would be more plausible if it were shown that performance in discrimination and differentiation procedures has the same quantitative properties. Consider one of them, the scalar property. Fetterman (1993) trained pigeons to discriminate two response numerosities n_1 and n_2 in the manner described above and then introduced probe trials with intermediate numerosities to obtain a psychometric function. This function plots the probability of choosing the "Large" alternative against the probe trial numerosity. Across experimental conditions, Fetterman varied the absolute values of n_1 and n_2 and their ratio. Results showed that when the ratio was held constant (e.g., 5 vs. 10, 10 vs. 20, and 20 vs. 40 for a ratio of 2; or 5 vs. 20, 10 vs. 40, and 20 vs. 80 for a ratio of 4), the psychometric functions were superposed when plotted with a relative-number scale, that is, when all the numerosities used in a condition were divided by the smaller numerosity of each pair. Consistent with the scalar property, the curves for different ratios were not superposed. The scalar property is a sort of Weber's law in the sense that equal ratios yield equal discriminabilities.

Whether the scalar property holds in numerosity differentiation studies is unclear. In Mechner's (1958) FCN schedule, the scalar property would be observed if the data obtained with different criteria also were superposed when plotted with a relative-number scale. To illustrate, assume that, when the reinforcement criterion required four responses on the left lever before a switch to the right lever yielded food, the rat produced a Gaussian distribution of responses with mean μ_1 and standard deviation σ_1. Assume similarly that, when the reinforcement criterion required eight responses, it produced a normal distribution with mean μ_2 and standard deviation σ_2. The scalar property requires that $\mu_2/\mu_1 = \sigma_2/\sigma_1$ or, equivalently, that the coefficient of variation or Weber fraction remain constant ($\sigma_1/\mu_1 = \sigma_2/\mu_2 = \gamma$). With one exception discussed below, no study using FCN schedules (e.g., Galbicka, Fowler, & Ritch, 1991; Laties, 1972; Mechner, 1958; Platt & Johnson, 1971) reported the data in ways that allow direct examination of the scalar property for a wide range of values of N (i.e., means and standard deviations were not presented).

The exception was Hobson and Newman's (1981) study with pigeons, which found decreasing coefficients of variation as N increased from 4 to 50. More specifically, the authors found that both μ and σ increased with N, but μ increased at a higher rate. Although important and clearly in need of replication, Hobson and Newman's findings are difficult to interpret because of a procedural feature. Instead of a pure FCN schedule, the authors used a mixture of a FCN schedule and a fixed-ratio (FR) schedule: On a random half of the trials, the pigeons received food after pecking the left key at least N times and then switching to the right key (FCN N); on the other half of the trials, the pigeons received food after the Nth peck on the left key (FR N). Consider then the differences in the contingencies of reinforcement that are implemented when

the two schedules are combined (in a mixed FCN–FR schedule) and their potential effects on the average and the variability of run length. Concerning the average run length, the prediction is straightforward: Because left-key pecks are reinforced directly only in mixed FCN–FR schedules, run length is probably greater under this schedule than under the equivalent FCN schedule. More generally, we predict that, when all other things are equal, the tendency to peck the left key is stronger when left key pecks are reinforced directly (the mixed schedule) than when they are reinforced indirectly, that is, when followed by a peck on the right key (the FCN schedule).

Concerning the variability of run length, the prediction is less clear. On the one hand, one could predict smaller variability in the mixed schedule because the pigeon has more opportunities to learn to concentrate its runs around N, the extra opportunities being the trials on which reinforcers follow the Nth peck on the left key. This prediction assumes that reinforcers after the Nth peck on the left key are functionally equivalent to reinforcers after runs of N pecks on the left key (i.e., after N pecks on the left key followed by one peck on the right key). On the other hand, with equal reasonableness one could predict greater variability in the mixed schedule because the variability of *reinforced* run lengths tends to be greater in that schedule. To illustrate with an extreme but easy-to-understand example, suppose a pigeon produces only runs of length 8 when the reinforcement criterion equals 4. If the schedule is FCN 4, then all reinforcers follow runs of eight pecks and the variability of reinforced run length is 0. But if the schedule is mixed FCN 4 FR 4, then half of the reinforcers follow runs of eight pecks and half follow runs of four pecks, and the variability of reinforced run lengths is positive. A pigeon sensitive to the reinforced run lengths would probably show greater run-length variability on the mixed FCN–FR schedule than on the FCN schedule.

Given the preceding arguments, we do not know whether Hobson and Newman's (1981) findings and conclusions concerning the scalar property extend to pure FCN schedules. Hence, the first goal of the present study was to clarify this issue. To that end, we extended Mechner's (1958) FCN schedule to pigeons and varied the numerosity criterion systematically across experimental conditions. We expected the results would yield the functions relating the average and standard deviation of run length to N and thereby show whether the coefficient of variation changed or remained constant.

Some researchers have argued that a (hypothetical) neural counting mechanism shares many components with a (hypothetical) neural timing mechanism. Meck and Church (1983; also Meck, Church, & Gibbon, 1985; Roberts, 1995) advanced the idea that there is only one mechanism and that it functions in either a timing mode or a counting mode. The difference is that, in the timing mode, the accumulator receives pulses from the pacemaker throughout the duration of the to-be-timed stimulus, whereas in the counting mode the accumulator receives pulses only for a brief period of time during each occurrence of the to-be-counted event. In fact, the similarities between timing and counting are not restricted to theory but extend to empirical issues. Most of the procedures used in numerosity discrimination and differentiation studies and the results obtained with them parallel procedures used in, and results obtained with, time discrimination and differentiation studies. For example, the bisection procedure is used to study both stimulus numerosity discrimination and time discrimination, and the results obtained with it, at least with animals, also are similar (i.e., the ogive-like psychometric functions have the point of subjective equality close to the geometric mean of the two training durations or numerosities; and the psychometric functions obtained with durations or numerosities with the same ratio overlap when plotted in relative time or number; see Gallistel, 1990; Meck & Church, 1983; and Meck, Church, & Gibbon, 1985).

The parallels between time and numerosity studies suggested the second goal of the present investigation, namely, to study whether pigeons can learn to emit not one but two response numerosities within a single experimental condition. The study is the numerosity differentiation analogue of the following study on temporal differentiation: A pigeon receives food for pecking a key after either 10 s or 120 s have elapsed since the onset of the trial. No cue signals whether the current trial will be short or long and the two types of trials are equally probable (a mixed fixed-interval 10 s

fixed-interval 120 s with its components randomly presented). The result of this training is that, during the long trials, the average rate of pecking increases from the beginning of the trial until approximately 10 s have elapsed, at which point the rate of pecking decreases before beginning to increase again until the end of the trial. That is, the response rate curve shows two peaks, one close to 10 s and the other close to 120 s. Moreover, the widths of the curve around the two peaks show the scalar property, for the width at 120 s is about 12 times the width at 10 s (Leak & Gibbon, 1995; see also Catania & Reynolds, 1968; Whitaker, Lowe, & Wearden, 2003).

In the present study we developed and tested a numerosity differentiation analogue of the preceding mixed schedule. In a modified FCN schedule, pigeons could receive food in one of two ways: (a) by pecking the left key exactly four times and then switching to the right key; or, if this did not occur (e.g., the pigeon pecked the left key five times and then switched); (b) by continuing to peck the left key for a cumulative total of at least 16 pecks and then switching to the right key. To put it differently and perhaps more intuitively, the pigeon had a first chance to obtain food by switching to the right key after emitting a run of four pecks on the left key. Intrinsic variability in its ''estimate'' of four (Galbicka et al., 1991; Laties, 1972; Mechner, 1958; Platt & Johnson, 1971) meant that on some trials the pigeon would switch to the right key after runs of, say, three or five pecks on the left key. On those trials, the pigeon had a second chance to obtain food by switching to the right key after emitting a second run of pecks on the left key. Food followed the second run provided the two runs on the left key added to at least 16 pecks. Thus, if the first run was three pecks long, then food was delivered if the second run was at least 13 pecks long. We predicted that the pigeons would generate distributions of run length on the left key with two peaks, one close to 4 and the other close to 16. Or, to put it differently, the pigeons would produce two runs, the first of about four pecks and the second of about 12 pecks such that the first and second runs combined added to 16 pecks. In addition, if the scalar property held within a single experimental condition, then the means and the standard deviations of the two peaks would be in approximately the same ratio.

Two experiments are reported below. The first experiment used the standard FCN schedule and varied the reinforcement criterion N across conditions. We examined two issues: how the mean μ and standard deviation σ of run length varied with N and whether the coefficient of variation σ/μ remained constant. The second experiment used a modified FCN schedule that shaped a bimodal distribution of run lengths. The issues under scrutiny were whether pigeons can learn to emit two response numerosities on the same key and within the same condition and how the ratios σ/μ for the two run-length distributions compare. Whereas Experiment 1 examined the learning of response numerosities successively (between conditions), Experiment 2 examined the learning of response numerosities simultaneously (within a condition). We conclude with some thoughts on how pigeons and rats learn to emit patterns with particular numbers of responses when exposed to FCN schedules.

EXPERIMENT 1

The experiment extended Mechner's (1958) FCN schedule to pigeons and varied the reinforcement criterion across conditions from 4 to 8 to 16 to 32. On the basis of previous studies using the same procedure (Galbicka et al., 1991; Laties, 1972; Mechner, 1958; Platt & Johnson, 1971), we predicted that in each condition the pigeons would generate Gaussian distributions of run length and that the mean and standard deviation of these distributions would increase linearly with the criterion N.

If the standard deviation and the mean increase linearly with N, then the coefficient of variation will either remain constant for all N or approach a constant as N increases. The former alternative corresponds to the strict version of Weber's law, that is, $\sigma = k \times \mu$ and therefore $\sigma/\mu = k$, for some constant $k > 0$ (the Weber fraction); the latter alternative corresponds to the generalized version of Weber's law, that is, $\sigma = k \times \mu + b$, and therefore $\sigma/\mu = k + b/\mu$. If we allow b to be greater than or equal to 0, then the latter alternative will include the former as a special case. The deviations from the scalar property will depend on the relative magnitudes of b and μ such that when b is much smaller than μ—as it

certainly is for large N and as it may be for small N—the term b/µ becomes negligible, the coefficient of variation approaches the constant k, and the distributions of run length superpose increasingly more when plotted by relative number.

The distributions obtained in this experiment and their estimated parameters served as a baseline against which the data from the second experiment could be compared.

METHOD

Subjects

Five pigeons (*Columba livia*) were maintained at 80 percent of their free-feeding body weight with water and grit continuously available in their home cages. Three pigeons (P99, P98 and P10) were 2 years old at the beginning of the study and had experience with variable-interval and variable-ratio reinforcement schedules. The other 2 pigeons (P16 and P13) were 1 year old and experimentally naïve.

Apparatus

The pigeons were studied in three identical operant chambers. Each chamber was 34 cm wide, 30.5 cm long, and 34 cm high. The walls and ceiling were made of aluminum, and the floor was wire mesh. The front panel was equipped with two circular keys that were 2.5 cm in diameter. The keys were centered on the wall, 23 cm above the floor, and 14 cm apart, center-to-center. The key on the right could be illuminated with green light and the key on the left with red light. On the back wall of the chamber, 3.5 cm below the ceiling, a houselight provided general illumination. Reinforcement consisted of mixed grain delivered through a food hopper. The hopper was accessible through a 6 cm × 5 cm opening that was centered on the front panel 8.5 cm above the floor. The operant chamber was enclosed in an outer box. On the back wall of the outer box, a ventilating fan circulated air through the chamber and helped to mask extraneous noises. A personal computer programmed in C++ controlled all experimental events and recorded the data.

Procedure

Preliminary training. The pigeons learned to peck the keys through autoshaping. Next they were exposed to trials in which, after a 10-s blackout intertrial interval (ITI), the houselight and the two keylights were illuminated, the left always with red and the right always with green light. The reinforcement rule was initially a mixed FCN FR schedule with a proportion p of FR trials and a proportion 1-p of FCN trials (Mechner, 1958). Thus, with probability p a reinforcer was delivered after the Nth peck on the left key (FR N schedule) and with probability 1-p a reinforcer was delivered after one peck on the right key provided that peck was preceded by at least N pecks on the left key (FCN schedule). In both cases the pigeon had to peck the left key a minimum of N times. If the pigeon pecked the right key before it completed N pecks on the left key, the trial was cancelled and the ITI started. Across five or six sessions of 60 trials each, parameter p decreased from .5 to 0 whereas parameter N increased from 1 to 16.

During the early sessions the reinforcement duration was adjusted for each pigeon in order to minimize extrasession feeding. Final values ranged from 2 to 4 s. During reinforcement, the houselight and the keylights were turned off, and the hopper light was illuminated.

Experiment proper. Sessions were divided into trials, and each trial was scheduled as follows: After the 10-s ITI, both keys were illuminated, the left with red and the right with green light. A peck on either key turned both keys off for 100 ms. After at least N consecutive pecks on the left key, the first peck on the right key was reinforced, and then the ITI started. A premature peck on the right key cancelled the trial by turning both keys off and starting the ITI. The trial also was cancelled if the pigeon did not peck for 120 s. Each session lasted until 100 trials were completed or 50 reinforcers were collected, whichever came first.

The experiment was divided into four conditions, each characterized by the value of the reinforcement criterion N. As Table 1 shows, the order of N was counterbalanced across pigeons according to a Latin square design with one repetition. Each condition remained in effect for a minimum of 10 sessions and until the mean and standard deviation of run length appeared visually stable for five consecutive sessions.

As Table 1 shows, for some pigeons N had to increase considerably between successive conditions (e.g., from N=4 to N=32). To prevent

Table 1

The reinforcement criterion N and the number of sessions (in parentheses) for each condition and pigeon. The asterisks indicate that the criterion increased gradually for some subjects between the two conditions and for the number of sessions shown in parentheses.

Pigeon	Condition 1		Condition 2		Condition 3		Condition 4
P16	32 (16)		4 (18)		16 (10)		8 (13)
P99	16 (15)	*(5)	32 (19)		8 (10)		4 (13)
P10	4 (16)		8 (18)	*(8)	32 (30)		16 (14)
P98	8 (16)		16 (25)		4 (23)	*(9)	32 (23)
P13	4 (20)		8 (27)	*(29)	32 (16)		16 (12)

extinction, in these cases (see asterisks in the Table) the criterion increased gradually across sessions. For example, for pigeon P98, the transition from Condition 3 (N=4) to Condition 4 (N=32) lasted nine sessions and during these sessions the criterion increased from 8 to 16 to 20 and to 24 before it was set at 32. Data analysis was based on the last five sessions of each condition.

RESULTS AND DISCUSSION

Figure 1 shows the relative frequency distributions of run length for each pigeon and reinforcement criterion. These distributions were well fit by Gaussian functions with two free parameters each, the mean μ and the standard deviation σ. Table 2 shows summary statistics obtained from the data and from the curve fits. In addition to μ and σ, the Table shows the coefficient of variation, $\gamma = \sigma/\mu$, and the variance accounted for, ω^2.

The run-length distribution was centered close to the reinforcement criterion. The difference between the obtained mean run length and N did not change consistently with N (repeated-measures ANOVA, $F(3,12) = 0.77$), but the relative difference (i.e., $(\mu-N)/N$) decreased with N (repeated-measures ANOVA, $F(3,12) = 3.76$, $p < .05$). The Gaussian functions accounted for 80 to 100 percent of the variance in the data (mean=95%).

The area of a distribution at and above the criterion corresponds to the proportion of reinforced trials. This area decreased with N: The averages for N=4, 8, 16, and 32 equaled 86%, 77%, 72%, and 54%, respectively (repeated measures ANOVA, $F(3, 12) = 6.7$, $p < .05$). This finding means that as N increased the pigeons became less efficient in earning food.

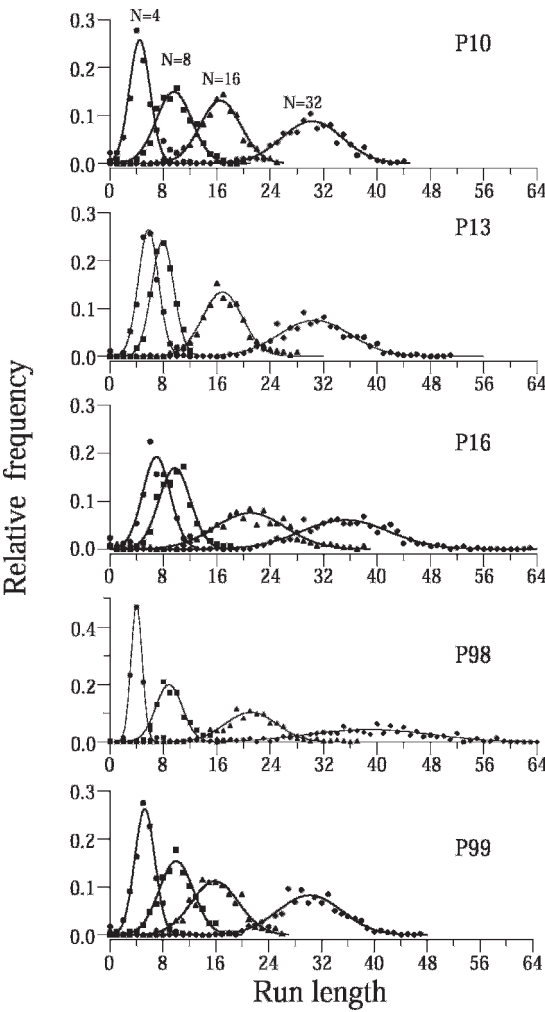Figure 2 shows how the two free parameters of the Gaussian functions as well as the

Fig. 1. The relative frequency distribution of run length during the last five sessions of each condition. A condition was defined by the reinforcement criterion N. The curves are the best-fitting normal distributions (see parameters in Table 2).

Table 2

Run-length statistics from curve fitting and from the data during the four conditions of Experiment 1. Statistics μ, σ, and γ are the mean, standard deviation, and coefficient of variation, respectively; ω^2 is the variance accounted for by the fitted curve.

		N = 4		N = 8		N = 16		N = 32	
		Fit	Data	Fit	Data	Fit	Data	Fit	Data
P10	μ	4.4	4.9	9.6	9.8	16.6	16.2	30.3	29.6
	σ	1.6	1.9	2.7	2.8	3.0	3.5	4.5	6.3
	γ	0.35	0.40	0.28	0.29	0.18	0.22	0.15	0.21
	ω^2	0.96		0.97		0.98		0.94	
P13	μ	5.7	5.9	7.9	7.9	16.9	17.2	30.6	30.8
	σ	1.5	1.8	1.7	1.6	3.0	3.3	5.3	5.6
	γ	0.26	0.31	0.21	0.21	0.18	0.19	0.17	0.18
	ω^2	0.98		0.99		0.97		0.94	
P16	μ	7.0	7.0	9.7	9.8	21.0	20.7	35.2	34.7
	σ	2.1	2.3	2.4	2.6	5.3	6.2	6.5	8.3
	γ	0.30	0.33	0.24	0.26	0.25	0.30	0.19	0.24
	ω^2	0.91		0.96		0.92		0.92	
P98	μ	4.0	4.2	8.8	9.2	21.1	21.2	38.9	37.1
	σ	0.8	1.1	2.0	2.3	3.9	4.4	9.2	11.1
	γ	0.21	0.26	0.23	0.25	0.18	0.21	0.24	0.30
	ω^2	1.00		0.97		0.96		0.80	
P99	μ	5.3	5.3	10.0	10.2	15.9	15.8	30.2	29.7
	σ	1.5	1.6	2.6	2.7	3.6	4.0	4.8	6.5
	γ	0.29	0.31	0.26	0.26	0.22	0.25	0.16	0.22
	ω^2	0.99		0.98		0.96		0.92	
Avg	μ	5.3	5.4	9.2	9.4	18.3	18.2	33.1	32.4
	σ	1.5	1.7	2.3	2.4	3.7	4.3	6.1	7.6
	γ	0.28	0.32	0.24	0.25	0.20	0.23	0.18	0.23
	ω^2	0.97		0.98		0.96		0.91	

coefficient of variation changed with N. The top panels show that for each pigeon the mean increased linearly with N. The slope of the regression lines varied from 0.87 to 1.26. (In fitting the data from pigeon P98 the intercept was forced to 0 because a negative intercept made no sense in the present situation.) The lines accounted for at least 99 percent of the variance. The middle panels show that the standard deviation also increased linearly with N. The slopes of the lines ranged from 0.1 to 0.3 (again, the intercept of the regression line for pigeon P98 was forced to zero). Variance accounted for ranged from 89 to 99 percent.

The bottom panels show the coefficient of variation calculated from the two estimated parameters (μ and σ) and the hyperbolas or straight line predicted from the equations for the mean (top panel) and standard deviation (middle panel) regression lines. Thus, from $\mu(N) = b_1 + m_1 N$ and $\sigma(N) = b_2 + m_2 N$, one gets for $\gamma(N)$ the hyperbola $(b_2 + m_2 N) / (b_1 + m_1 N)$, which converges to the asymptotic value m_2/m_1. For pigeon P98, because both b_1 and b_2 equaled 0 the coefficient of variation equaled m_2/m_1 for all N.

Figure 3 summarizes the behavior of the mean, standard deviation, and coefficient of variation as a function of N for all subjects. In each panel, the filled circles show the mean value of the parameter across pigeons (± 1 standard deviation), and the lines show the mean of the individual functions displayed in Figure 2. The top panel shows the linear increase of the mean run length with N. The line has a slope of 0.99 and an intercept of 1.65 and accounts for more than 99 percent of the variance. The middle panel shows the linear increase of the standard deviation of run length with N. The best-fitting line has a slope of 0.16 and an intercept of 1.05 and accounts for more than 99 percent of the variance. Finally, the bottom panel shows the coefficient of variation decreasing and approaching the value of 0.16, the ratio of the slopes of the two regression lines. The fitted curve accounts for 97 percent of the variance.
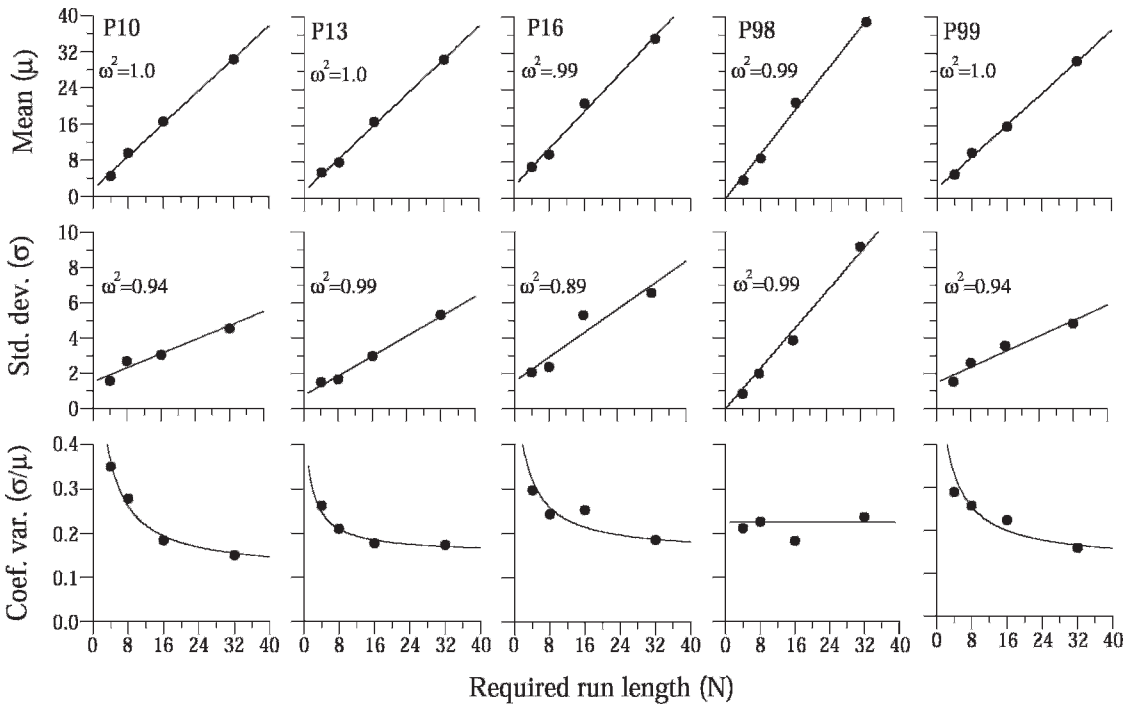
Fig. 2.   The estimated means (top panels) and standard deviations (middle panels) of run length as a function of the reinforcement criterion N; the bottom panels show the coefficients of variation. The solid lines are the best-fitting regression lines.

To determine how the preceding results compare with previous findings, Figure 4 shows the data from the present study together with the data from four other studies—Mechner (1958), Platt and Johnson (1971), Laties (1972), and Hobson and Newman (1981). Consider the first two studies. Both used rats and varied N from 4 to 8 to 12 to 16. Mechner[1] showed the histogram of run length for each of his 6 animals and for each criterion value N. Platt and Johnson showed histograms for only 1 rat but mentioned that the selected histograms represented well the data from the other 7 animals. We fitted a normal density function to each histogram and thus estimated the mean and the standard deviation of the distribution of run length for each criterion N. For Mechner's study, we then averaged the estimated parameters across the 6 rats. As the top two panels of Figure 4 show, the rat data were similar to the present data—the two parameters increase linearly with N, and the

best-fitting regression lines relating the estimated parameters to N had similar slopes: For the mean, the average slopes were 1.15 (Mechner), 1.18 (Platt & Johnson) and 0.99 (present study); for the standard deviation, the average slopes were 0.17 (Mechner), 0.19 (Platt & Johnson), and 0.16 (present study).

The solid diamonds in Figure 4 show the averages of the observed (not estimated) means and standard deviations reported by Laties (1972) for 4 pigeons at criterion value N = 8. The observed values were close to the values obtained in the present experiment.

Consider now Hobson and Newman's (1981) data displayed in the bottom two panels. As mentioned before, their pigeons were exposed to mixed FCN–FR schedules in which the required run length N varied from 4 to 50. The authors report that the group average of the mean run lengths was well fitted by the power function $\mu(N) = 1.6 \times N^{0.92}$ ($\omega^2 = 0.999$). For that reason, instead of reading the data points from a figure, we used the power function to generate them. The results (see the third panel) show that, in Hobson and Newman's experiment, μ increased at a slightly

---

[1] Mechner's (1958) data contain a mistake in Figure 1, rat N4, as the proportions do not add to 1 but to .5. We multiplied all values by 2.
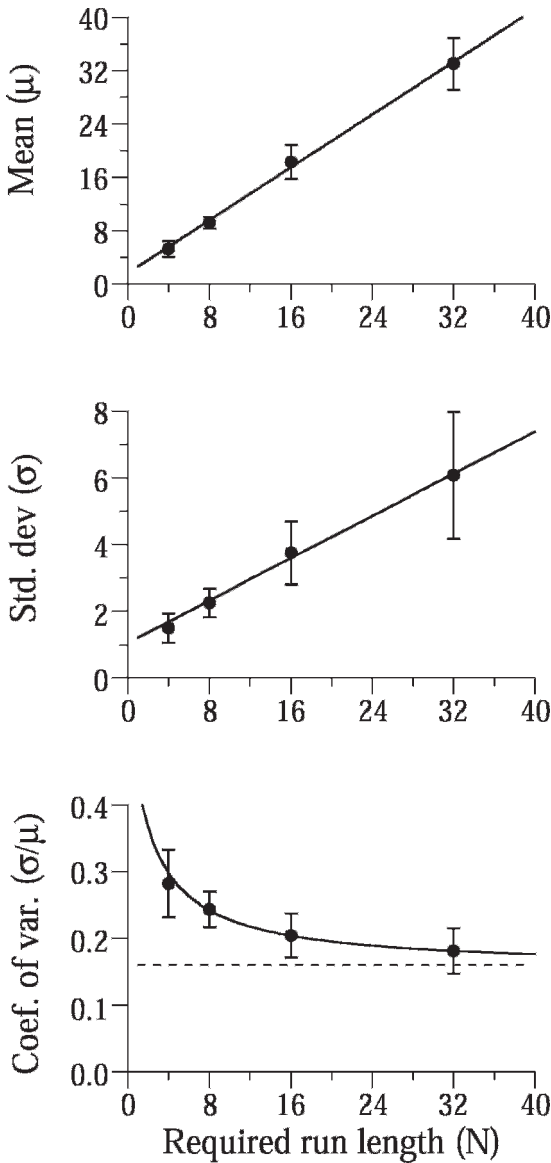
Fig. 3. The means across pigeons of the estimated means (top), standard deviations (middle) and coefficients of variations (bottom) as a function of the reinforcement criterion N. The bars show ± one standard deviation. The solid lines are the averages of the best-fitting individual lines shown in Figure 2. In the bottom panel, the dotted line shows the asymptote approached by the mean coefficient of variation.

faster rate than in the present experiment: The best-fitting regression lines had equations $\mu(N) = 1.14 \times N+1.72$ and $\mu(N) = 1.0 \times N+1.53$, respectively, and accounted for more than 99 percent of the variance. The direct reinforcement of left key pecks in Hobson and Newman's study may have been the cause of the slightly longer runs produced by their subjects.

The bottom panel shows the coefficients of variation and the best-fitting hyperbolas for Hobson and Newman's (1981) data and the current data. Hobson and Newman's data were read from Figure 8-4 in their paper. To fit the hyperbola to these data, we used the equation reported above for $\mu(N)$ (i.e., $1.14 \times N+1.72$) and estimated the slope and intercept parameters of the corresponding equation for the standard deviation, $\sigma(N) = m \times N+b$. The best-fitting values were $m = 0.22$ and $b = 1.6$, which means that the hyperbola displayed in the panel has equation $\gamma(N) = (0.22 \times N+1.6)/(1.14 \times N+1.72)$. The hyperbola for the present experiment is $\gamma(N) = (0.16 \times N+0.96)/(1.0 \times N+1.53)$. The two curves accounted for more than 98 percent of the variance. The difference in the slopes and intercepts of the standard deviation functions indicate that run length was more variable in the mixed FCN–FR schedule than in the pure FCN schedule. As mentioned above, the cause of this difference may have been the higher variability of reinforced runs in the mixed schedule.

The data in the bottom panel also show that in both studies the group average of the coefficient of variation decreases with N rapidly and then approaches an asymptote. Although not shown in the panel, the same plots with Mechner's (1958) and Platt and Johnson's (1971) data sets reveal a similar trend—a rapid decrease of the mean coefficient of variation with N such that for values of N greater than about 10 (Mechner and Platt & Johnson with rats) or 12 (Hobson & Newman and the present study with pigeons) the mean coefficient of variation does not change appreciably. It follows that the scalar property is clearly violated for small values of N but holds increasingly better for values of N greater than about 12.

In conclusion, the results from Experiment 1 show that, in FCN schedules, the distributions of run length are roughly Gaussian. The mean and the standard deviation of these distributions increase linearly with the reinforcement criterion, N. A generalized version of Weber's law holds in the sense that, as N increases, the coefficient of variation decreases hyperbolically and approaches a con-
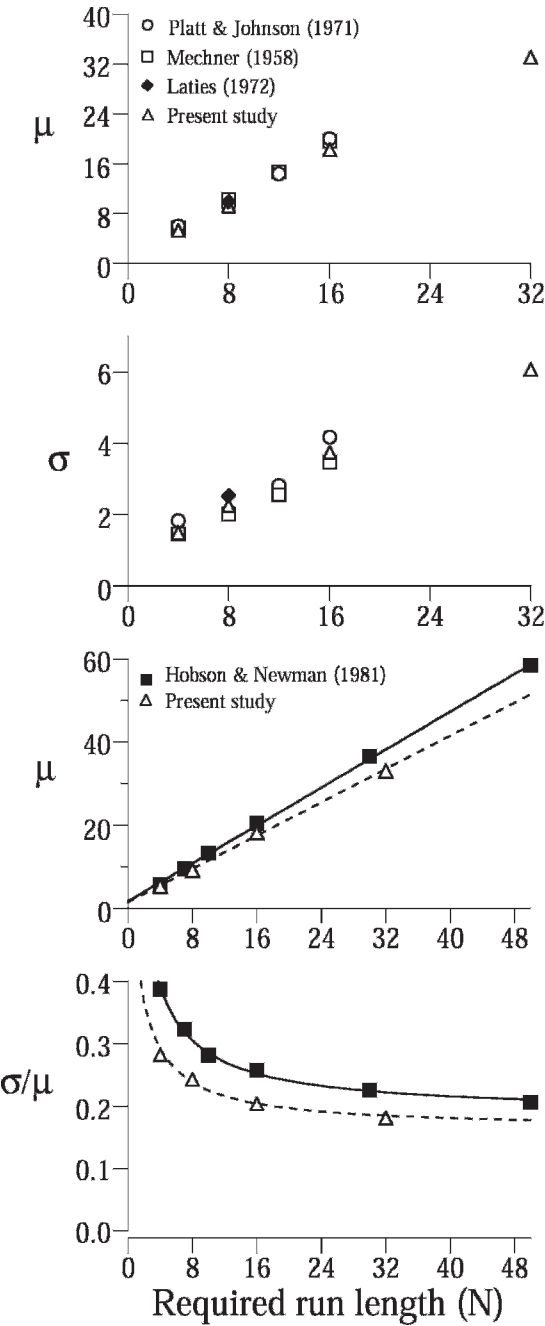
Fig. 4. Group data from the present study (open triangles) are compared with the results from other studies. Mechner (1958) used 6 rats in a chamber with two levers. The open squares show the averages of the means (μ) and standard deviations (σ) estimated to fit individual frequency distributions. Platt and Johnson (1971) used 8 rats in a chamber with a lever (response A) and a hole for nose poking (response B). The open circles show the means and standard deviations estimated when fitting the data of a single rat. Laties (1972) used 4

stant value. These results are consistent with previous findings obtained with FCN schedules using either rats or pigeons. Whether these findings hold also when two numerosities are learned within the same experimental session was addressed in the second experiment.

EXPERIMENT 2

The reinforcement rule of Experiment 2 was the numerosity equivalent of the mixed fixed-interval fixed-interval (*mixed* FI FI) schedule used by Catania and Reynolds (1968). In that study, pigeons received food at the end of a fixed interval that with probability p was 30 s long and with probability $1-p$ was 240 s long. No stimulus signaled which FI schedule would be in effect on each trial. The authors varied p across experimental conditions to study how that parameter affected the response-rate function. In the present experiment, the pigeons received food for producing two numerosities, one equal to 4 and the other equal to 16 or more. Similar to Catania and Reynolds' experiment, the smaller numerosity was reinforced with probability p, and parameter p varied across conditions; the larger numerosity was reinforced always with probability 1.0.

The main goal of the experiment was to examine whether pigeons could learn two response numerosities simultaneously. We predicted they would generate two runs of responses on the left key. The first run would have a mean length of about 4 and correspond to the effect of the reinforcers received for pecking the left key exactly 4 times and then switching. The second run would have a mean length slightly greater than 12 and correspond to the effect of the reinforcers received for pecking the left key at least 16 times and then switching. To state these predictions more precisely, let r_1 and r_2 stand for the lengths of the first and second runs, respectively. That is, r_1 is the number of pecks emitted on the left

←

pigeons with two response keys. The filled diamond shows the averages of the obtained means and standard deviations. Hobson and Newman (1981) used 12 pigeons with mixed FR–FCN schedules. Their data appear as filled squares in the lower two panels that display results for the means (μ) and the coefficient of variation (σ/μ).

key from the trial onset until the pigeon switches to the right key, and r_2 is the number of pecks emitted on the left key from the end of the first run until the pigeon switches again to the right key. In the following sequence of left and right pecks, LLRRLLLLLR, the first run is two pecks long ($r_1 = 2$) and the second run is five pecks long ($r_2 = 5$). Given the reinforcement contingencies of Experiment 2, we predicted that: (a) the distribution of r_1 would be Gaussian with a mean close to 4; (b) the distribution of r_2 would also be Gaussian with a mean slightly greater than 12; (c) given (a) and (b), the mean of $r_1 + r_2$ would be slightly greater than 16; (d) given the hyperbola obtained in Experiment 1 (see bottom panel of Figure 3) the coefficient of variation of r_1 would be slightly greater than the coefficient of variation of r_2; and (e) the frequency of r_1 would decrease with p, that is, the strength of short runs would decrease as their probability of reinforcement decreased.

In Experiment 2, when a pigeon completed a short first run and did not receive food it could emit a second run. Because reinforcement followed whenever the two runs added up to at least 16, one could expect, in a well-trained pigeon, a negative correlation between r_1 and r_2: short first runs followed by long second runs and, conversely, long first runs followed by short second runs. Such correlations were reported by Platt and Johnson (1971; see also Gallistel, 1990) when their rats were exposed to FCN schedules in which premature switches from response A (lever press) to response B (nose poke) did not cancel the trial. When there is no penalty for switching before the criterion N is met, the animals can produce on each trial a minimum of one run ($r_1 \geq N$) and a maximum of N runs (by alternating strictly between A and B, i.e., $r_1 = 1$). Platt and Johnson showed that the rats tended to produce a relatively long first run followed by (when $r_1 < N$) much shorter runs. In addition, they found negative correlations ranging from –.16 to –.39 between the current run length r_i and the cumulative number of lever presses since trial onset (Σr_j, j = 1..i–1). They concluded that the cumulative number of presses since trial onset was a discriminative stimulus controlling run length after premature switches to B.

In the present experiment the pigeons could switch from the left key to the right key only twice and therefore only two runs could be produced on each trial. The procedure with a maximum of two runs may be seen as intermediate between a regular FCN schedule in which only one run is allowed and Platt and Johnson's (1971) procedure in which N runs are allowed. If the length of the second run is controlled by the length of the first run, then r_1 and r_2 should be negatively correlated. Alternatively, if r_1 has no influence on r_2, then their correlation should be zero. Experiment 2 also addressed this issue.

METHOD

Subjects and Apparatus

Four experimentally naive pigeons (*Columba livia*) participated in the study. Housing conditions and the experimental apparatus were the same as in Experiment 1.

Procedure

The experimental details differed from Experiment 1 only in the reinforcement rule. In Experiment 1 all trials ended after the first run with reinforcement if $r_1 \geq N$ and with a time-out if $r_1 < N$; the criterion N varied across conditions. In Experiment 2 each trial ended after the first or the second run. If $r_1 \geq 16$, reinforcement followed always and the trial ended. If $r_1 = 4$ then reinforcement followed with probability p. If reinforcement did occur, then the trial ended, but if reinforcement did not occur, then the pigeon could produce a second run. The second run ended in reinforcement provided $(r_1 + r_2) \geq 16$; otherwise the ITI followed immediately. Thus the pigeon could earn a reinforcer in one of three ways: (a) by emitting a first run at least 16 pecks long; (b) by emitting a first run exactly four pecks long; and (c) by emitting two runs that summed to 16 or more pecks.

During four successive experimental conditions, the probability of reinforcement given that $r_1 = 4$ equaled 1.0, .5, .25, and 1.0. Table 3 shows the details. For example, consider pigeon P11. During Condition 1, this pigeon was exposed to the reinforcement schedule described above with $p = 1.0$. That is, reinforcement followed the first run provided $r_1 = 4$ or $r_1 \geq 16$. In case $r_1 \neq 4$ and $r_1 < 16$, reinforcement followed the second run provided $r_1 + r_2 \geq 16$. In all other cases the trial ended without food. During Condition 2, p changed to .5, which meant that only 50

Table 3

Schedule parameters and number of sessions for each condition of Experiment 2. Reinforcement occurred with probability p if the first run was of length N_1 and with probability 1.0 if the first run was of length at least N_2 or the two runs combined were of length at least N_2. Asterisks indicate sessions with shaping (see text for further details).

Pigeon	Condition	N_1	N_2	p	Sessions
P11	1	4	16	1.0	43
	2	4	16	.5	16
	3	4	16	.25	23
	4	4	16	1.0	22
P12	1	4	16	1.0	28
	1*	4	8	1.0	11
	1*	4	12	1.0	8
	1	4	16	1.0	32
	2	4	16	.5	24
	3	4	16	.25	24
	4	4	16	1.0	19
P9	1	4	16	1.0	28
	1*	4	8	1.0	20
	1*	–	4		7
	1*	4	12	1.0	7
	1*	–	4		18
	1	4	16	1.0	43
	2	4	16	.5	25
	3	4	16	.25	17
	4	4	16	1.0	22
P18	1	4	16	1.0	30
	1*	4	8	1.0	16
	1	4	16	1.0	32
	2	4	16	.5	23
	3	4	16	.25	29
	4	4	16	1.0	19

percent of the trials with $r_1 = 4$ were followed by food. It was still the case that $r_1 \geq 16$ guaranteed food or that, when the first run was not reinforced, the pigeon could receive food by producing a second run such that $r_1 + r_2 \geq 16$. Condition 3 was similar to Condition 2 with p equal to .25, and Condition 4 was equal to Condition 1.

When first exposed to Condition 1, only pigeon P11 learned that reinforcement followed short runs of $r_1 = 4$ pecks. As will be shown below, the pigeon produced distributions of r_1 and r_2 close to 4 and 12, respectively. The other 3 pigeons showed no such learning; after about 30 sessions they continued to produce distributions with no appreciable frequencies around 4. A shaping-like procedure then was implemented (as indicated by the asterisks in Table 3). Consider the conditions for pigeon P12, for example. Reinforcement followed when $r_1 = 4$, $r_1 \geq 8$, or $r_1 + r_2 \geq 8$. By decreasing N_2 from 16 to 8,

we hoped to increase the frequency of runs of length 4 and thereby increase the pigeon's contact with reinforcement at $r_1 = 4$. When the distribution of first runs showed appreciable frequencies close to 4, we increased N_2 from 8 to 12 and then from 12 to 16. For pigeon P9, it was necessary to eliminate N_1 altogether during a few sessions in order to increase the frequency of short runs. Without N_1 the reinforcement contingency is identical to that of Experiment 1, a pure FCN schedule with $N = 4$. Using this shaping-like procedure the 3 pigeons learned the reinforcement contingencies as evidenced by a bimodal distribution of run length. Each experimental condition remained in effect until the means and standard deviations of run length appeared visually stable for five consecutive sessions.

RESULTS AND DISCUSSION

The data analysis covered the last five sessions of each condition. Figure 5 shows the distributions of the response runs during Condition 1. The left panels correspond to r_1, the middle panels to r_2, and the right panels to the two runs combined (i.e., r_1 and $r_1 + r_2$). Because the number of opportunities to produce the two runs differed markedly, the Figure shows absolute frequencies. The curves in the left and middle panels are the best-fitting Gaussian density functions. The curves in the right panels will be described later. Table 4 shows the estimated mean μ and standard deviation σ for each curve and the variance accounted for, ω^2.

In the left panels, the modes of r_1 were at 3 or 4, the means were slightly less than 4 (the mean for all pigeons $<\mu> = 3.4$), and the standard deviations were slightly greater than 1 ($<\sigma> = 1.2$). Note that, although first runs of 16 or more pecks were reinforced always, the pigeons rarely produced them during the last five sessions of Condition 1. As Table 4 shows, the curves fitted the data well ($<\omega^2> = .97$).

In the middle panels, the modes of r_2 ranged from 11 to 15, the means were close to 14 ($<\mu> = 14.3$), and the standard deviations close to 4 ($<\sigma> = 4.0$). The curves accounted for a smaller proportion of the variance than those for the first run ($<\omega^2> = .87$). The differences between the means of r_1 and r_2 show that all pigeons differentiated the two consecutive run lengths.
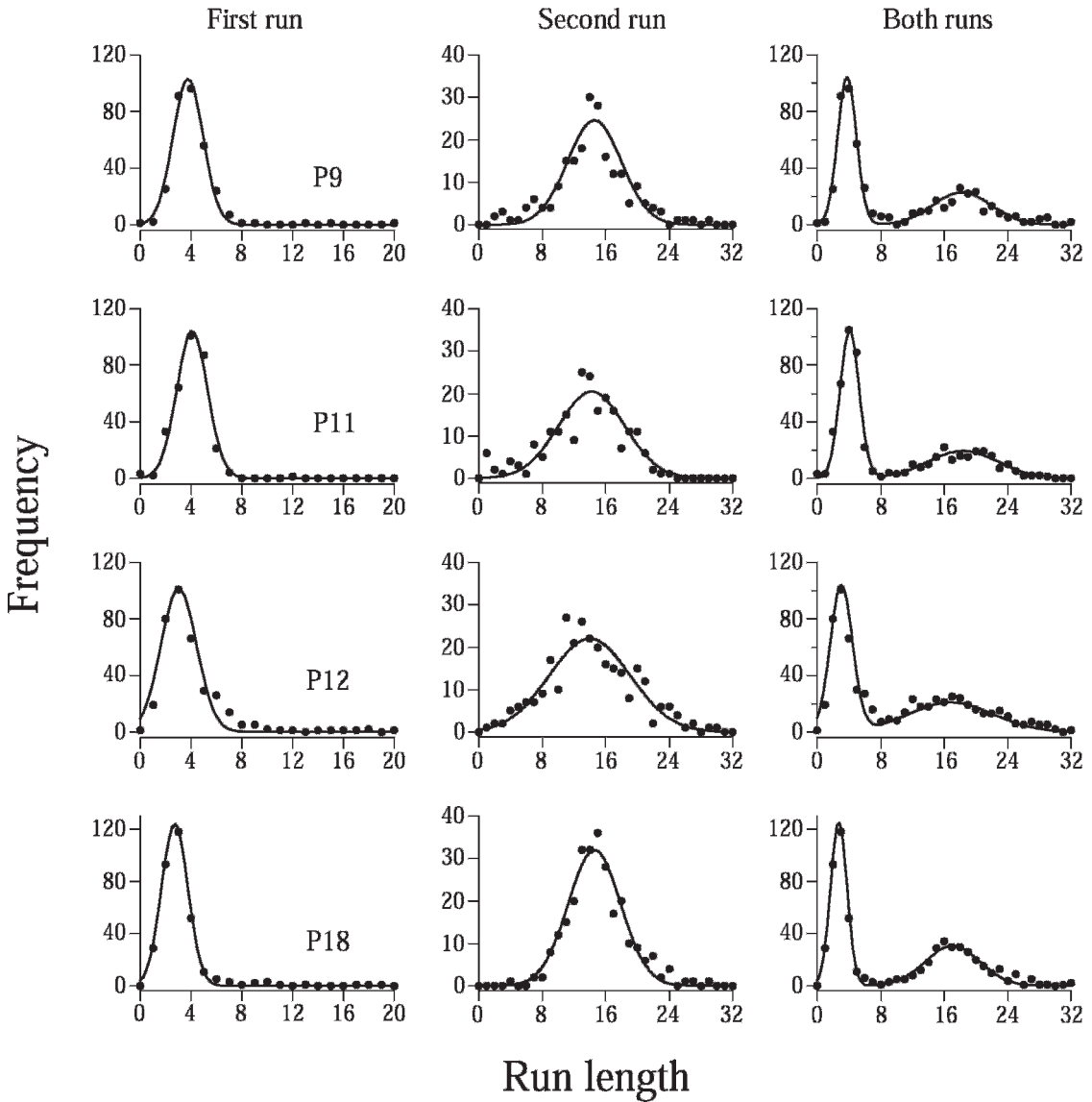
Fig. 5. The absolute frequency distributions of the first run (r_1, left panels), second run (r_2, middle panels), and both runs (r_1 and $r_1 + r_2$, right panels) during Condition 1 of Experiment 2 ($p = 1.0$). The curves in the left and middle panels are best-fitting normal distributions with parameters (μ_1, σ_1) and (μ_2, σ_2), respectively (see Table 4). The curves in the right panels plot Equation 1 using the same parameters as the left and middle curves (i.e., μ_1, σ_1, μ_2, and σ_2).

The right panels combine the distributions of r_1 (left peaks) and $r_1 + r_2$ (right peaks). To understand how these distributions were obtained and what they mean, imagine that a counter is reset to zero at the beginning of each trial and then increments by one whenever the pigeon pecks the left key. However, the counter is visible only when the pigeon switches to the right key. If one records the number visible on the counter whenever

the pigeon switches to the right and then plots the frequency distribution of the recorded numbers, one gets the distributions shown in the right panels. The importance of these distributions is that they show the run lengths on the left key that presumably cued the pigeon to check the right key for food. In all cases, the pigeons checked for food either after approximately 3 or 4 pecks (r_1) or after slightly more than 16 pecks ($r_1 + r_2$) on the left

Table 4

Run-length statistics based on curve fitting in Experiment 2. Except for pigeons P9 and P11 in Condition $p = .25$, parameters μ_1 and σ_1 were used to fit the distribution of first runs (r_1) and parameters μ_2 and σ_2 were used to fit the distribution of second runs (r_2). The three ω^2 values are the variance accounted for by the curves on the left, middle, and right panels, respectively, in Figures 5, 6, 7, and 8. For pigeons P9 and P11, parameters μ_1, σ_1, μ_2, σ_2, and λ (see text for explanation) were used to fit the r_1 data in Figure 8. No attempt was made to fit the data of Pigeon 11 in the middle and right panels of Figure 8 (cells with ''—'').

Condition	Pigeon	μ_1	σ_1	μ_2	σ_2	λ	ω_1^2	ω_2^2	ω_3^2
$p = 1.0$ first	P9	3.8	1.2	14.6	3.4		0.98	0.85	0.97
	P11	4.1	1.2	14.2	4.2		0.98	0.79	0.98
	P12	3.0	1.4	14.0	5.1		0.94	0.89	0.93
	P18	2.7	1.0	14.6	3.3		0.99	0.96	0.99
$p = 1.0$ last	P9	3.6	1.2	12.2	4.5		0.98	0.88	0.96
	P11	3.7	0.7	14.2	4.6		1.00	0.63	0.99
	P12	3.3	1.5	10.5	3.3		0.91	0.90	0.87
	P18	2.8	1.1	16.2	3.9		0.99	0.93	0.97
$p = .5$	P9	3.5	1.3	13.6	4.6		0.99	0.82	0.97
	P11	3.7	0.9	12.6	3.5		1.00	0.86	1.00
	P12	3.6	1.8	11.8	3.5		0.92	0.91	0.90
	P18	3.5	1.3	14.1	4.5		0.92	0.86	0.89
$p = .25$	P9	3.9	1.1	11.1	3.8	0.27	0.79	0.92	0.87
	P11	4.5	1.0	12.1	5.7	0.23	0.90	—	—
	P12	13.4	4.4	4.8	3.3		0.94	0.88	0.95
	P18	2.9	1.3	14.8	3.6		0.94	0.92	0.93

key. Obviously these numbers were due to the reinforcement contingencies for r_1 and r_2.

The Appendix shows how to derive the equation for the curves that appear in the right panels. Assuming independence between the two runs (an assumption to which we will return), the probability of a run length equal to n, $P(R = n)$, is given by:

$$P(R = n) = f(n, \mu_1, \sigma_1)$$

$$+ \sum_{i=0}^{\min(n-1, 15)} \{f(i, \mu_1, \sigma_1) f(n - i, \mu_2, \sigma_2)\}$$

$$- f(4, \mu_1, \sigma_1) f(n - 4, \mu_2, \sigma_2) p, \qquad (1)$$

where $f(i, \mu, \sigma)$ is the normal density function with mean μ and standard deviation σ evaluated at run length i.

The curves in the right panels of Figure 5 show Equation 1 with the parameters used to fit the data from r_1 (μ_1 and σ_1; see curves in the left panels) and r_2 (μ_2 and σ_2; see curves in the middle panels). In other words, no extra parameters were used in the curve fitting. The curves fit the data well ($<\omega^2> = .97$).

Figure 6 shows the data from the last condition in which p again equaled 1.0. With a few exceptions, the results reproduced those from Condition 1: The distributions of r_1

remained centered at 3 or 4 and those of r_2 at values between 10 and 16. The exceptions were some excessively short or excessively long runs of r_2 (see the arrows in the middle panels, at $r_2 = 1$ for Pigeon P11, and at $r_2 \geq 22$ for Pigeon P9). The curves fit the data as well as in Condition 1: $<\omega^2>$ equaled .97, .84, and .95 for the left, middle, and right panels, respectively.

Figure 7 shows the data for Condition 2. When only 50 percent of the $r_1 = 4$ runs were reinforced, the distributions of r_1 and r_2 did not change appreciably. Thus r_1 remained centered at 3 or 4 pecks and r_2 at values between 12 and 14. For Pigeon P12, the frequency of $r_1 = 0$ runs increased substantially (see arrow and compare that value with the corresponding values in Figures 5 and 6.). This result suggests that, when the probability of reinforcement decreased, stimulus control of left key pecking at trial onset also decreased. Loss of stimulus control also is suggested by the relatively high frequency of short second runs (i.e., $1 \leq r_2 \leq 4$; see arrow). The curves fit the data as well as in Conditions 1 and 4: $<\omega^2>$ equaled .96, .86, and .94 for the left, middle, and right panels, respectively.

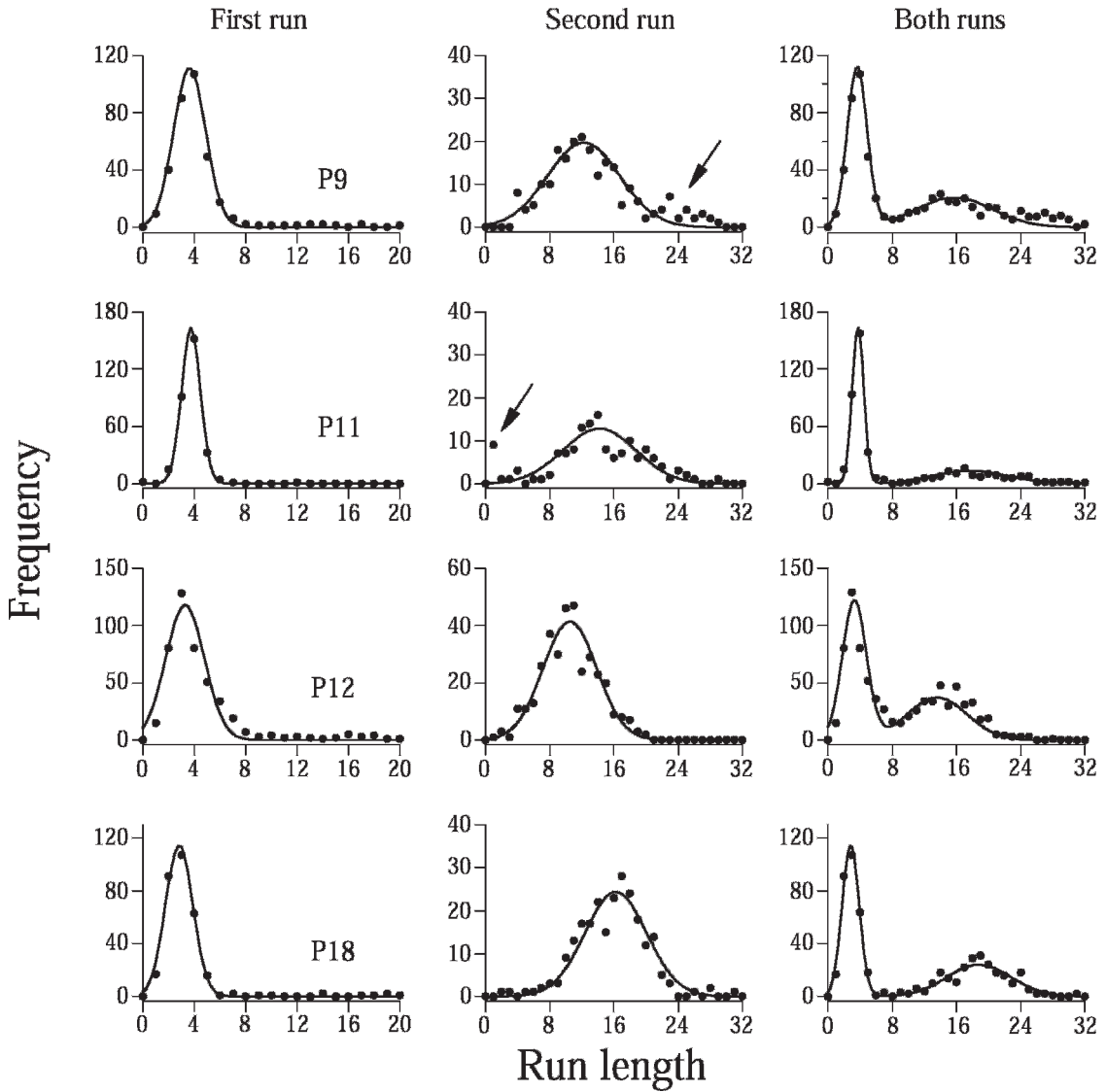Finally, Figure 8 shows the data from Condition 3. When the probability of reinforce-

Fig. 6. The absolute frequency distributions of the first run (r_1, left panels), second run (r_2, middle panels), and of both runs (r_1 and $r_1 + r_2$, right panels) during Condition 4 of Experiment 2 ($p = 1.0$). The curves in the left and middle panels are best-fitting normal distributions with parameters (μ_1, σ_1) and (μ_2, σ_2), respectively (see Table 4). The curves in the right panels plot Equation 1 using the same parameters as the left and middle curves (i.e., μ_1, σ_1, μ_2, and σ_2). The arrows show excessively short or excessively long runs.

ment for r_1 = 4 decreased to .25, the distributions of r_1 and r_2 did not change for pigeon P18 but changed significantly for the other 3 pigeons, although in different ways. Consider pigeon P12. The initial runs of 3 or 4 pecks almost disappeared and were replaced by much longer runs with 13.5 pecks on the average. The second runs were much shorter than in the other conditions and averaged 4.8 pecks. That is, compared with the other

conditions, the average lengths of r_1 and r_2 were reversed—instead of emitting a short and then a long run, the pigeon emitted a long run and then, if it did not receive food, a short run. A Gaussian density function fit reasonably well the distribution of r_1 ($\omega^2 = 0.94$) but less well the distribution of r_2 ($\omega^2 = .88$). Equation 1 fit well the combination of both runs ($\omega^2 = 0.95$).

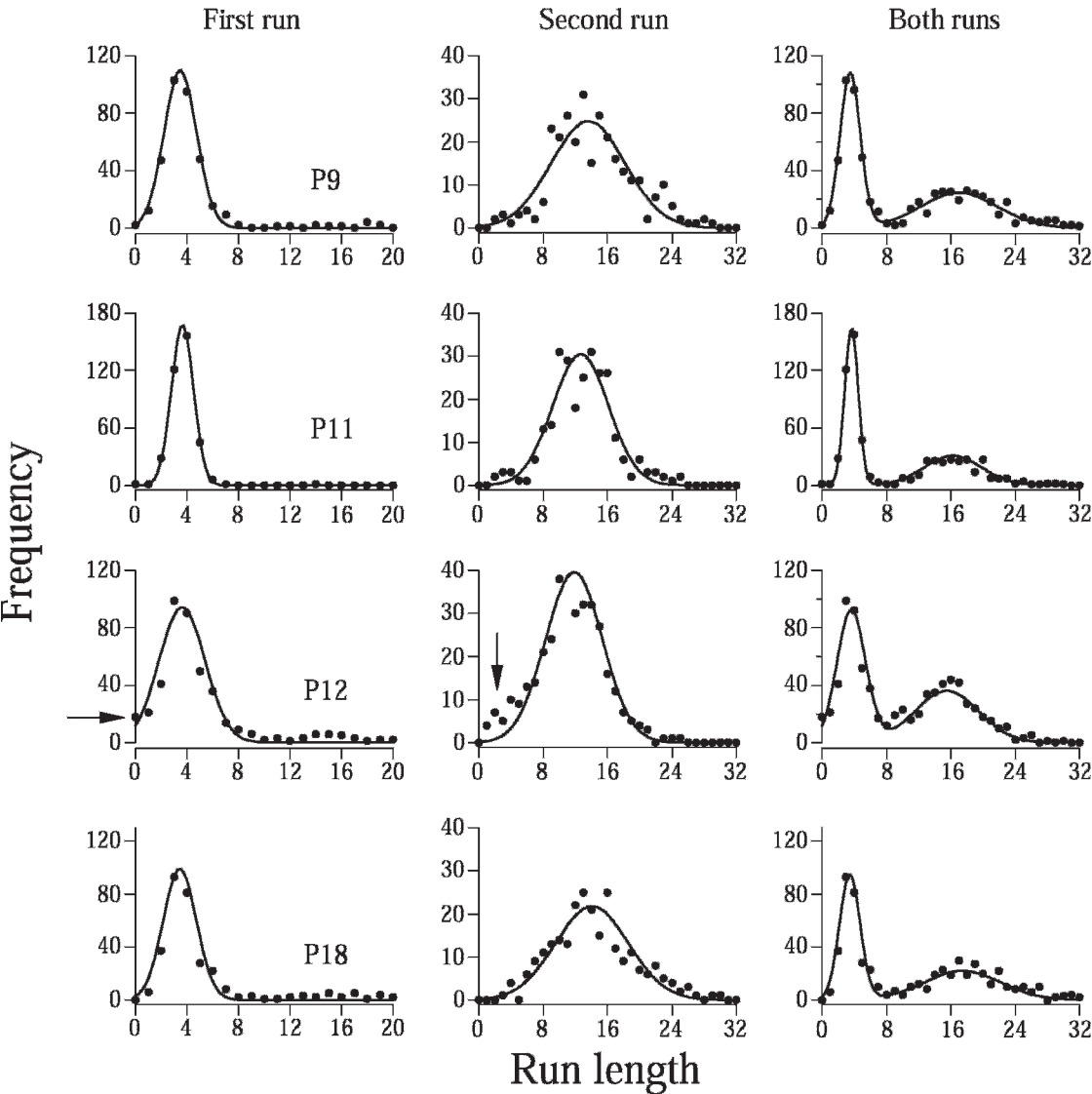Next consider pigeon P9. The distribution of r_1 had three modes: one at 0, another at 4,

Fig. 7. The absolute frequency distributions of the first run (r_1, left panels), second run (r_2, middle panels), and of both runs (r_1 and $r_1 + r_2$, right panels) during Condition 2 of Experiment 2 ($p = .5$). The curves in the left and middle panels are best-fitting normal distributions with parameters (μ_1, σ_1) and (μ_2, σ_2), respectively (see Table 4). The curves in the right panels plot Equation 1 using the same parameters as the left and middle curves (i.e., μ_1, σ_1, μ_2, and σ_2). The arrows show excessively short runs.

and another at 16. The distribution of r_2 remained similar to the distributions observed during the other conditions. The relatively high frequency of runs of length 0 (see arrow) suggests a loss of stimulus control at trial onset. We ignored such runs when fitting the data. The distribution of the remaining runs could be fit with the following model: At trial onset the pigeon sampled from a normal distribu-

tion f, with parameters similar to those obtained during the other conditions ($\mu_1 = 3.9$, $\sigma_1 = 1.1$). However, at the end of the run the pigeon did not always peck the right key. Instead, on a fraction of the trials the pigeon emitted a second run by sampling from another normal distribution with parameters $\mu_2 = 11.2$ and $\sigma_2 = 3.8$. In other words, we suggest that the two response patterns learned
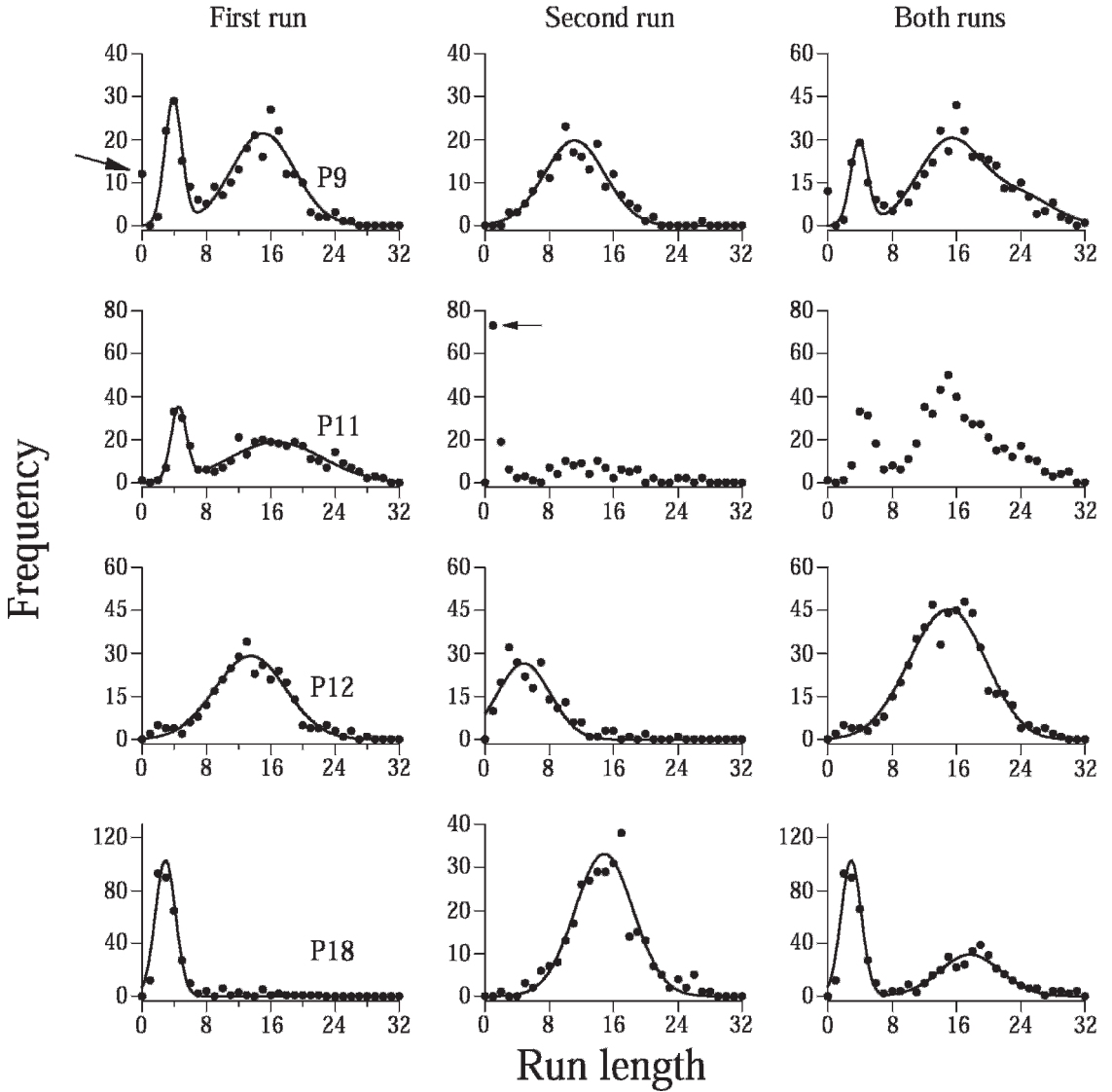
Fig. 8. The absolute frequency distributions of the first run (r_1, left panels), second run (r_2, middle panels), and of both runs (r_1 and $r_1 + r_2$, right panels) during Condition 4 of Experiment 2 ($p = .25$). Left panels: For pigeons P12 and P18, the curves are best-fitting normal distributions with parameters (μ_1, σ_1). For pigeons P9 and P11 the curves are weighted averages of two normal distributions (see Equation 2), one with weight λ and the other with weight $1-\lambda$ (see Table 4). Middle panels: For pigeons P9, P12, and P18 the curves are best-fitting normal distributions with parameters (μ_2, σ_2). Right panels: For pigeons P9, P12, and P18 the curves plot Equation 1 using the same parameters as the left and middle curves. The arrows show excessively short runs.

during the previous conditions, namely, "r_1 pecks on the left key followed by one peck on the right key" (pattern 1) and, in case the trial continued, "r_2 pecks on the left key followed by one peck on the right key" (pattern 2), may have been disrupted such that on some trials the pigeon jumped, as it were, from the first to the second pattern *without pecking the right key*.

If this interpretation is correct, then the first run would be fitted by a weighted average of two distributions:

$$P(r_1 = n) = \lambda \times f(n,\mu_1,\sigma_1) + (1 - \lambda)$$
$$\times f\left(n,\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2} \right), \quad (2)$$

where λ represents the proportion of trials in which the pigeon pecked the right key at the end of the first run. In case the first run was not reinforced, the pigeon emitted a second run by sampling from the same normal distribution with parameters μ_2 and σ_2. The solid curves in the top panels show that this model fit the data reasonably well (see ω^2 values in Table 4).

For pigeon P11, the distribution of r_1 also was well fitted by a weighted average of two normal density functions obtained as follows: At trial onset the pigeon sampled from a normal distribution with parameters $\mu_1 = 4.5$ and $\sigma_1 = 1.0$. With probability $\lambda = .23$ it ended the run by pecking the right key; with the complementary probability $(1-\lambda = .77)$ it did not peck the right key but continued the trial by sampling from another normal distribution with parameters $\mu_2 = 12.1$ and $\sigma_2 = 5.7$ and then pecking the right key. However, in contrast with pigeon P9, the distribution of second runs displayed in the middle panel of Figure 8 was not Gaussian-like, for it presented a high frequency of very short runs (i.e., $r_2 = 1$ or 2; see arrow) and a low frequency of longer runs ranging from 8 to about 18. That is, the distribution of r_2 also seemed to be a mixture of two distributions. Unlike the previous cases, no simple model consisting of a weighted average of normal density curves fit well all the data of pigeon P11. This fact explains the absence of theoretical curves in the middle and right panels of Figure 8.

In summary, when the reinforcement probability for runs of four pecks decreased to $p = .25$, the pigeons generated distinctly different distributions of r_1. For pigeon P18 the distribution remained centered at short runs; for pigeon P12 it was centered at long runs; and for pigeons P9 and P11 it was bimodal, resembling the combined distributions of r_1 and r_1+r_2 observed during the other conditions. The distributions of r_2 also fell into three classes. For pigeons P18 and P9 it remained centered at long runs; for pigeon P12 it was centered at short runs; and for pigeon P11 it defied any simple description.

Concerning the coefficients of variation of the r_1 and r_2 distributions, γ_1 and γ_2, respectively, two questions may be asked: Were they significantly different? And, were they consistent with the values obtained in Exper-
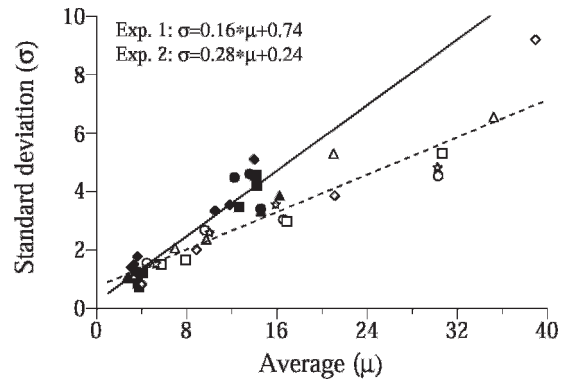

Fig. 9. The estimated standard deviations plotted against the estimated means of the run-length distributions in Experiment 1 (empty symbols) and Experiment 2, Conditions $p = 1.0$ and $p = .5$ (filled symbols). Each symbol corresponds to a different pigeon. The two lines are the average of the individual best-fitting regression lines and account for 88 percent of the variance in the data of Experiment 1 and 89 percent in the data of Experiment 2.

iment 1? To answer the questions we analyzed the data from the $p = 1.0$ and $p = .5$ conditions and excluded the data from the $p = .25$ because of the marked changes in run-length distributions observed during that condition.

On the basis of the results from Experiment 1, in particular the hyperbolic decay of γ with N and therefore also with μ, one would expect γ_2 to be slightly less than γ_1. The data were in the expected direction: In Condition 1 ($p = 1.0$), the average values were $\gamma_1 = 0.36$ and $\gamma_2 = 0.28$; in Condition 2 ($p = .5$) the average values were 0.37 and 0.31, and in Condition 4 ($p = 1.0$) they were 0.34 and 0.31. The overall averages were $\gamma_1 = 0.36$ and $\gamma_2 = 0.30$. However, due to the variation across pigeons and their small number, the differences between γ_1 and γ_2 did not reach statistical significance. A 3-condition × 2-distribution repeated measures ANOVA yielded all ps > .05.

In Experiment 2 the average coefficient of variation for r_1 ($\gamma_1 = 0.36$) was close to the value expected on the basis of Experiment 1 (0.33 for a mean run length of 3.4), but the average coefficient of variation for r_2 ($\gamma_2 = 0.30$) was clearly above the value expected on the basis of Experiment 1 (0.21 for a mean run length of 13.5). To better understand these results, Figure 9 plots the estimated standard deviations against the estimated means of the

individual run length distributions (Table 2 may be consulted to obtain the values of σ and μ in Experiment 1, and Table 4 may be consulted to obtain the values of σ_1, μ_1, σ_2, and μ_2 in Experiment 2). For the two data sets to be consistent, the filled and empty symbols should fall along the dotted line in the Figure. This happened for the run lengths of 3 or 4 pecks that characterized r_1 (leftmost filled circles) but not for the run lengths of 11 to 15 pecks that characterized r_2 (rightmost filled circles). It follows that σ increased with μ more in Experiment 2 than in Experiment 1. A *t*-test comparing the slopes of the individual best-fitting regression lines yielded a significant result $[(t(7) = 3.5, p < .05)]$.

The difference in the slopes of the lines may stem from a behavioral or a statistical source (or both). Concerning the behavioral source, note that in Experiment 1 all runs were preceded by the ITI, whereas in Experiment 2 the first runs were preceded by the ITI, but the second runs were preceded by one or more pecks on the right key. This difference in the stimulus conditions at the onset of some runs may explain the differences in run-length variability. Concerning the statistical source, if runs follow a normal distribution with the standard deviation a linear function of the mean (the generalized Weber's law), then a negative correlation between r_1 and r_2 is sufficient to increase the variability of the second run. Specifically, if the conditional distribution of r_2 given r_1 equals a normal density function with parameters:

$$\mu_{r_2|r_1} = a - b \times r_1, \text{ and}$$

$$\sigma_{r_2|r_1} = c + d \times \mu_{r_2|r_1},$$

then it can be shown that the variance of r_2 will be given by:

$$\sigma_{r2}^2 = (c + d\mu_{r2})^2 + b^2\sigma_{r1}^2(d^2 + 1),$$

where μ_{r2} is the expected value of r_2 and σ_{r1} is the standard deviation of r_1. The expression shows that the variance of r_2 is greater than $(c+d \times \mu_{r2})^2$, which is the variance expected in the absence of a dependency of r_2 on r_1.

The preceding analysis raises the issue of how the length of r_1 correlated with the length of r_2. Table 5 shows the Pearson correlation coefficients for all pigeons and conditions. Half of the correlations were close to zero and

Table 5

Pearson coefficients of correlation between first and second runs in the four conditions of Experiment 2. Bold values are significantly different from 0 (*t*-test, $p < .05$).

Pigeon	$p = 1.0$ (first)	$p = 1.0$ (last)	$p = .5$	$p = .25$
P9	0.06	0.01	0.01	**−0.34**
P11	**0.15**	0.00	0.11	**−0.75**
P12	**−0.24**	**−0.19**	**−0.36**	**−0.29**
P18	−0.05	−0.08	**−0.19**	**−0.17**

statistically not significant (*t*-test, $p > .05$). The other half were statistically significant and, with one exception, negative. Pigeon P12 showed consistently significant negative correlations, and all pigeons showed significant negative correlations when $p = .25$. In fact, the data suggest a trend towards stronger negative correlations as p decreased. However, with the exception of one case, analyzed below, all coefficients were rather small; in fact, r_1 accounted for less than 12 percent of the variance of r_2. The magnitudes of the correlation coefficients were similar to those reported by Platt and Johnson (1971). The smallness of most correlation coefficients explains why the data displayed in the rightmost panels of Figures 5 to 8 could be fitted by Equation 1, which assumes independence between r_1 and r_2. In fact when, instead of Equation 1, we used an equation that took into account the potential negative correlation between r_1 and r_2, the quality of the fits did not change appreciably (i.e., ω^2 did not change by more than $0.01)^2$.

The only substantial correlation between r_1 and r_2 ($r = -.75$) occurred in pigeon P11 when $p = .25$. A close analysis of the scatter plot of the two run lengths showed that the correlation was due to fact that first runs shorter than about eight pecks tended to be followed by second runs with an average length of 13 pecks—as in the other experimental conditions—but first runs longer than eight pecks tended to be followed by second runs of one or two pecks.

---

[2] Instead of Equation 1, we fitted an equation that assumed that r_2 depended linearly, but with negative slope, on r_1. Specifically, the probability that $r_2 = j$ given that $r_1 = i$ equaled the Gaussian function f(i, μ(i), σ(i)) with mean μ(i) = a – b x i and standard deviation σ(i) = γ x μ(i), for positive constants a, b, and γ.

GENERAL DISCUSSION

Experiment 1 extended Mechner's (1958) FCN schedule to pigeons and varied the reinforcement criterion over a relatively large range ($4 \leq N \leq 32$). Pigeons received food for pecking key A at least N times and then switching to key B; premature switches to key B cancelled the trial. The results showed that pigeons can learn to switch to key B after emitting an average number of responses close to the reinforcement criterion. As the average number of pecks on key A increases, absolute precision (measured by the standard deviation of run length) increases linearly and therefore relative precision (measured by the coefficient of variation or Weber fraction) decreases hyperbolically to an asymptote. In addition, the distribution of run lengths is well fit by a Gaussian function.

These findings, together with similar findings reported by previous investigators using FCN schedules (see Galbicka et al, 1991; Gallistel, 1990; Laties, 1972; Mechner, 1958; Platt & Johnson, 1971), corroborate the idea that response numerosity is a differentiable dimension of behavior consistent with a generalized form of Weber's law—the scalar property was clearly violated for small numerosities but held for relatively large values of N. In addition, these findings add to the literature a systematic set of run-length distributions and their statistics (means and standard deviations) that cover a relatively large range of the numerosity dimension. As we show below, these distributions and statistics are important for the development and testing of theoretical models of numerosity differentiation.

Since Mechner's (1958) original study, researchers have asked if performance under FCN schedules is a case of temporal, not numerosity, differentiation. The question is raised because, if animals respond at a roughly constant rate (as our pigeons did), then run time will be strongly correlated with run length—hence, performance could be based on time, number, or both. Although we cannot rule out this possibility, two indirect arguments suggest that number was the important dimension. First, Mechner and Gevrekian (1962) found that postreinforcement pauses and run durations could be changed by manipulating deprivation level, but those changes did not affect the distribu-tions of run length for FCN schedules. In other words, deprivation affected only the nonreinforced dimension. Second, after a review of the literature from temporal and numerical procedures, Hobson and Newman (1981) concluded that the decreasing trend of the coefficient of variation observed in numerosity differentiation tasks for relatively small N (see the bottom panels of Figures 2 and 3) typically does not occur for comparable time-based procedures, which tend to yield constant coefficients of variation. The authors also concluded from their review of transfer tests (from time-based to number-based schedules or vice-versa) that subjects tend to be most sensitive to the reinforced dimension.

Experiment 2 attempted to differentiate two response numerosities within the same session. To that end, pigeons were exposed to a modified FCN schedule in which they had two ways to receive food. First, as in the standard FCN schedule, they could peck the left key at least 16 times and then switch to the right key, in which case reinforcement followed with probability 1.0. And second, they could gamble, so to speak, and try to collect food earlier and with less effort by pecking exactly four times on the left key and then switching to the right key. In this case food followed with probability p. If the pigeon decided to gamble but did not receive food, then it had a second chance of obtaining food by returning to the left key and pecking it for a cumulative total of at least 16 pecks. Parameter p varied across conditions from 1.0 to .25.

The results of Experiment 2 show that pigeons can learn two response numerosities simultaneously. This conclusion is based on the fact that, when p equaled 1.0 or .5, each pigeon emitted two response runs with non-overlapping numerical distributions: The first run had an average length close to 3.5 and a standard deviation close to 1.2, and it corresponded to the pigeons' attempts to earn the food available for pecking the left key exactly four times. The second run had an average length close to 13.5 and a standard deviation close to 4.0. It corresponded to the pigeons' attempts to earn the food available for pecking the left key at least 16 times (both runs were included in the count).

The preceding findings extend the results obtained by Fetterman, Dreyfus, and Stubbs (1985). In their study, pigeons received food

by pecking a key n_1 times (e.g., 8) provided the key was illuminated with red light, or by pecking the same key a total of n_2 times (e.g., 32) provided it was illuminated with green light. The pigeon could change the keylight color from red to green by pecking another key once. The schedule is similar to Hobson and Newman's (1981) mixed FR–FCN schedule in the sense that, in both cases, optimal performance required the pigeon to peck the main key n_1 times and then, if reinforcement did not occur, switch and peck the other key. Fetterman et al. reported that pigeons performed as follows: They pecked on the red key a number of times and then, when food was not presented, they changed over to green and continued to peck until food was presented or the trial ended. The distribution of the number of pecks on red before the changeover occurred—the equivalent of run length in FCN schedules—was Gaussian-like with a median close to the harmonic mean of n_1 and n_2.

It is important to clarify what Fetterman et al.'s (1985) results show and do not show. The distribution of run length on the red key shows that pigeons can learn to differentiate one response numerosity, and the median at the harmonic mean of n_1 and n_2 shows that differentiation is influenced by the schedule parameters. But the results do not show that pigeons can learn to differentiate two response numerosities. In fact, their experiment required and obtained only one numerosity and for that reason the authors could report only one run-length distribution for each pair of schedule parameters. To engender the differentiation of two response numerosities and thereby be able to report two run-length distributions (similar to those displayed in the left and middle panels of Figures 5 to 8), the authors would have to require a second changeover after a total of n_2 pecks on green. But then their experiment would be functionally equivalent to our Experiment 2 except for the keylight color changes on the main key. Whether these changes would alter our major empirical findings remains to be determined.

Experiment 2 also examined the effects of the probability of reinforcement following runs of four pecks (p) and the correlations between r_1 and r_2 (r). By analogy with mixed FI FI schedules, we expected that the relative weights of the r_1 and r_2 distributions would change directly with p or, more specifically, that the proportion of short runs (r_1) would decrease as p decreased. Results did not confirm this expectation: When p changed from 1.0 to .5 the r_1 and r_2 distributions did not change appreciably (compare Figures 5, 6, and 7), and when p changed from .5 to .25 these distributions changed appreciably in 3 out of 4 pigeons but in different ways (compare Figures 7 and 8). As for the correlations between r_1 and r_2, about half were not significantly different from 0; the other half were mostly negative but, as in Platt and Johnson's (1971) work, still weak ($r < 0.36$). Only one correlation was relatively strong ($r = -.75$ for pigeon P11 in Condition 3). More generally, there was some evidence that the correlations between r_1 and r_2 become more negative as p decreases (see Table 5).

Our results also suggest that the stimulus conditions at the onset of a run may influence its numerical precision. In fact, when $p = 1.0$ and $p = .5$, the standard deviations of the second runs tended to be greater than the value expected on the basis of the results of Experiment 1 (see Figure 9). We suggest that this difference may have been due to the (weak) negative correlations between r_1 and r_2 or to the fact that the stimulus conditions at the onset of the runs differed (ITI versus pecks on the right key). But how could the stimulus conditions at the onset of a run affect its variability? One way would be by affecting the number of responses required to initiate the hypothetical internal counter. That is, in the generalized form of Weber's law, $\sigma(N) = c \times N + d$, the context could determine the value of the constant term d such that the ITI would be associated with a smaller d than a peck on the right key. The hypothesis remains to be tested.

Toward a model of numerosity differentiation. The results of the two experiments may be interpreted in terms of acquisition of response patterns or runs of responses on key A followed by a switch to key B. The contingencies of reinforcement, in general, and the total number of responses on key A when reinforcement is delivered, in particular, shape the mean length, or numerosity, of the response runs. But how does the animal learn a pattern defined by response numerosity? How does response differentiation take place along the dimension of number? These general questions are related to three specific sub-questions.

First, under FCN schedules, all patterns whose length is equal to or greater than N are reinforced whereas all patterns whose length is less than N are extinguished. Therefore, reinforcement selection is directional in the sense that the mean length of the reinforced runs (say, μ^+) is at least as large as the mean length of all emitted runs (μ), that is, $\mu^+ > \mu$ (this condition will hold unless all run lengths are greater than N)[3]. Hence why wasn't there an upward drift of the mean run length on key A?

Second, and related to the previous point, the data from the N = 16 and N = 32 conditions in Experiment 1 show that pigeons can produce relatively long response runs. Why then are these runs not produced during the N = 8 or N = 4 conditions? If they were, the pigeons would collect many more of the available reinforcers. To illustrate, in Experiment 1 if the individual distributions obtained when N = 16 had been produced when N = 8, then the average percentage of reinforced trials would not equal 77 (the range across pigeons was 60 – 85) but 98 (with a range of 97 – 100; see also Galbicka et al., 1991, for another instance in which the mean of the normal-like distributions was significantly below the criterion).

And third, how does a pigeon learn that the reinforcement criterion decreased? To see the problem, consider a pigeon exposed to an FCN 32 schedule until it emits response runs with 32 pecks on the average. Then the schedule changes to FCN 4 as it did for pigeon P16 in Experiment 1. The result is that, although long response runs continue to be reinforced during the new condition, mean run length decreases to values close to 4. How did the pigeon detect the change in the criterion? (The opposite change, from 4 to 32, is easy to detect because of extinction following the change.) The problem is functionally equivalent to that of detecting extinction in an avoidance procedure (i.e., the removal of the aversive stimulus) when the animal is successfully avoiding the stimulus.

The preceding problems have not been addressed in previous discussions of models of counting as they apply to FCN schedules. As

noted in the Introduction, perhaps the best-known counting model is Meck and Church's (1983) adaptation of Scalar Expectancy Theory. According to this information-processing model, the number of pulses emitted by a pacemaker following each response (the counted event in FCN schedules) is added in the accumulator. When a switch to key B is reinforced, the number of pulses in the accumulator is multiplied by a random variable (the main source of the scalar property) and then transferred to memory. To decide on each trial when to switch to key B, the animal extracts a sample from the memory store and then compares it with the current number in the accumulator. The similarity between the two numbers determines whether the animal switches or not. The problem is that, without further assumptions, the model will be unstable. Because of the directional selection mentioned above, the mean of the distribution of counts in long-term memory should drift upward, which is contrary to fact[4]. Moreover, the model remains silent on *how* the animal detects changes in the criterion, for example, from 32 to 4. In what follows we propose a model that addresses the foregoing issues.

The model focuses on learning and takes for granted the scalar property. It assumes that at the onset of trial n the animal samples from a Gaussian distribution with mean $\mu(n)$ and standard deviation $\sigma(n)$, with $\sigma(n)$ equal to $\gamma \times \mu(n)$ (Weber's law). The number sampled on trial n, X, is the run length on that trial. Reinforcement will follow provided X $\geq$ N. According to the model, learning consists of changing μ as a function of trial outcome. Figure 10 shows the details.

There are three types of trials: (a) non-reinforced trials (case X_1 in the Figure); (b) reinforced trials in which the current run length is greater than the mean run length (case X_2); and (c) reinforced trials in which the current run length is less than the mean

---

[3] For an alternative, nondirectional or centripetal reinforcement selection, a schedule in which extinction occurs on both sides of the criterion, see Galbicka, Fowler, & Ritch (1991).

[4] Because in the application of Scalar Expectancy Theory to numerosity differentiation tasks the contents of the accumulator are stored in memory only during reinforced trials, it can be shown that the mean of the stored values will drift upwards. Without additional assumptions, this instability problem will occur not only in FCN schedules but also in the analogous Differential Reinforcement of Low Rates (DRL) schedules. To our knowledge this instability problem has not received the attention it deserves.
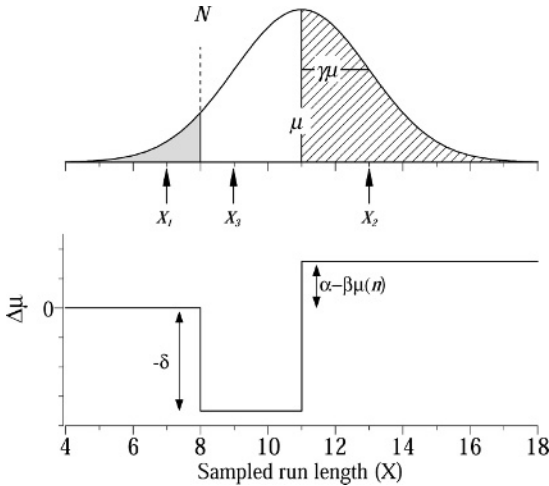
Fig. 10. A model of numerosity differentiation under FCN schedules with parameter N. On trial n, the animal samples a number X from a normal distribution with mean $\mu(n)$ and standard deviation $\sigma(n) = \gamma \times \mu(n)$. If X, the run length on that trial, is less then N, then the trial ends in extinction and $\Delta\mu(n) = 0$. If $X \geq N$ and $X < \mu(n)$, the trial ends in reinforcement and $\Delta\mu(n) = -\delta$. If $X \geq N$ and $X \geq \mu(N)$, the trial ends in reinforcement and $\Delta\mu(n) = \alpha - \beta\mu(n)$. The following relations hold: $\alpha > 0$, $\beta > 0$, $\delta > 0$, and $\delta > \alpha$.

run length (case X_3). The model assumes that parameter μ changes differently in each case.

(a) If the run is not reinforced ($X < N$), then the current mean run length does not change ($\Delta\mu = 0$). This assumption may have to be relaxed in more complex models and after experiments clarify how μ changes in extinction.

(b) If the run is reinforced and its length is greater than μ ($X \geq N$ and $X \geq \mu$), then μ increases by the amount $\Delta\mu = \alpha - \beta \times \mu(n)$, with $\alpha > 0$ and $\beta > 0$. That is, the magnitude of the increment decreases linearly with μ. This assumption captures the idea that, as μ increases, each reinforcer becomes more costly and therefore less effective in changing behavior. The assumption also ensures that μ will not increase beyond the value α/β (when $\mu > \alpha/\beta$, $\Delta\mu < 0$). The value α/β is the maximum mean run length the animal will sustain in the current circumstances (deprivation level, reinforcement quality and amount, etc.). One important consequence follows: When the maximum mean run is less than the reinforcement
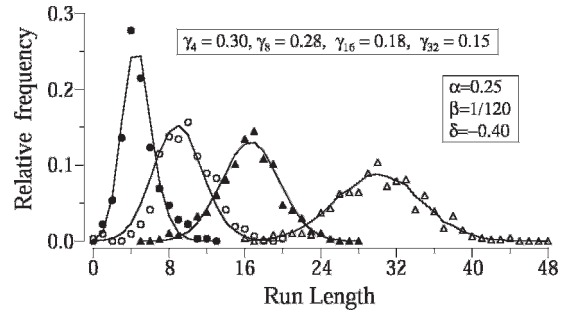

Fig. 11. The numerosity differentiation model fitted to the data of pigeon P10 (Experiment 1). Only the coefficient of variation γ was allowed to vary with N. The curves are averages of 100 simulations for each value of N.

criterion (i.e., $\alpha/\beta < N$), the model generates run-length distributions with means less than N (see Figure 1, N = 32, pigeons P10, P13, and P99). But when $\alpha/\beta > N$, the model generates distributions with means greater than N (the remaining cases in Figure 1). Thus the model captures one key property of the data.

(c) If the current run is reinforced but its length is less than μ ($X \geq N$ and $X < \mu$), then μ will decrease by a constant amount $\Delta\mu = -\delta$, with $\delta > 0$ and $\delta > \alpha$. The inequality $\delta > \alpha$ means that decreasing μ is easier than increasing it.

In summary, the three equations for $\Delta\mu$ describe how reinforcement shapes μ in the direction of N: Extinction trials do not change μ; reinforced trials may increase, decrease, or not change μ depending on how the reinforced run length compares with current μ and with the maximum sustainable μ.

Figure 11 shows the model fitted to the data of pigeon P10 from Experiment 1. The learning parameters α, β, and δ were kept constant, and only the coefficient of variation γ was allowed to vary with N, as the data suggest it should. The curves represent the average of 100 simulations with the initial μ close to N. The model fit the data reasonably well (.87 $\leq \omega^2 \leq$ .96). In particular, note that, in this case, the average run length is below the criterion because $\alpha/\beta < N$. For the remaining pigeons, the individual fits accounted for an average of 80 percent of the variance (range: 42% to 94%).

Figure 12 summarizes the fit to all data from Experiment 1. To compare model and data,
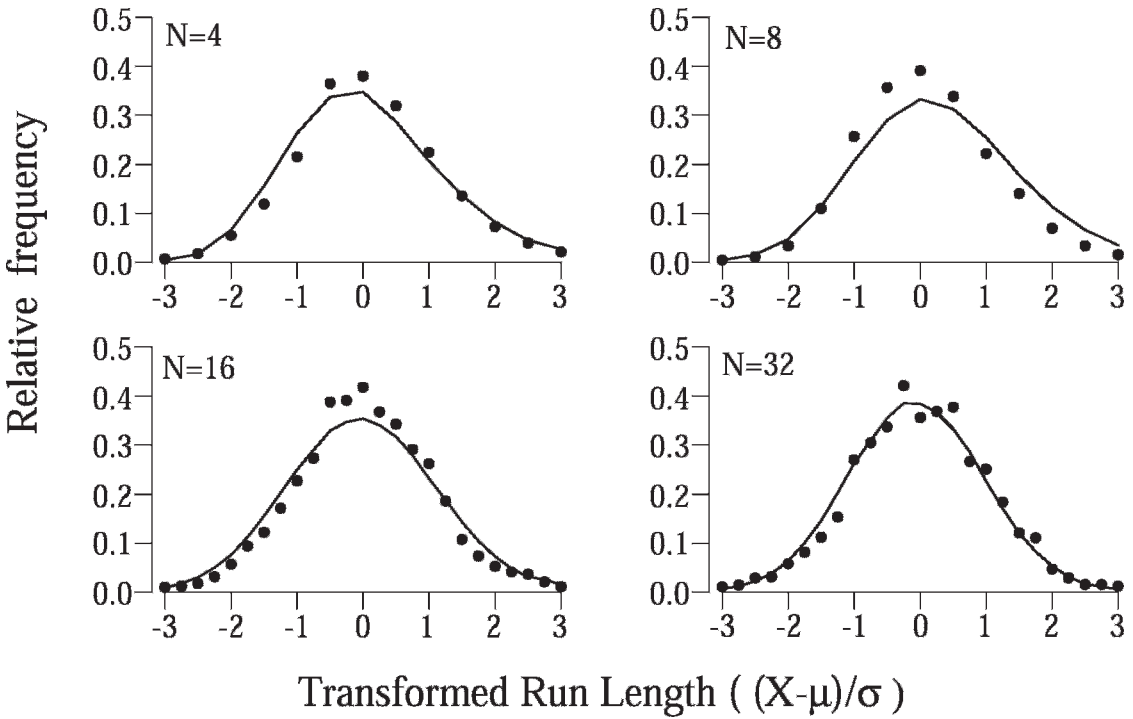
Fig. 12.   The symbols show the averages of the individual frequency distributions for the 5 pigeons of Experiment 1. The curves show the averages of the individual curves predicted by the numerosity differentiation model.

we first transformed the data from each pigeon and condition by subtracting a constant from each run length and dividing the result by another constant. The two constants were not the average and standard deviation of the data—which would have yielded z-scores—but the average and standard deviation of the Gaussian curves displayed in Figure 1. We refer to these transformed values as ''z-values.'' The same two constants were used to transform the curve predicted by the model. Next we used linear interpolation to obtain the ''data'' at regular z-values (from $z = -3.0$ to $z = +3.0$), and finally we averaged across pigeons to obtain the average data for each experimental condition. The same steps were carried out for each curve generated by the model, yielding the curves shown in Figure 12.

The average of the predicted curves fit the average of the obtained curves reasonably well both in terms of accuracy (mean) and precision (standard deviation), but they tended to undershoot the peak values observed around $N = 4$, 8, and 16. The proportion of variance accounted for equaled .98 ($N = 4$), .95 ($N = 8$), .97 ($N = 16$), and .98 ($N = 32$). Note that

the theoretical model used seven parameters to fit four curves whereas the atheoretical fits displayed in Figure 1 used eight parameters.

The model solves the three problems mentioned above. First, directional selection will not cause an upward drift in the mean run length for two reasons: because some reinforcers will actually decrease mean run length (whenever $X < \mu$), and because even when the reinforcers increase mean run length (whenever $X > \mu$) their effectiveness decreases with μ and vanishes when $\mu = \alpha/\beta$. Second, pigeons do not become more efficient at collecting the reinforcers because, whenever mean run length is much larger than the schedule criterion ($\mu \gg N$), the net effect of reinforcement is to decrease μ (because $\delta > \alpha$) and thereby reduce the subject's efficiency. Third, for the same reason, pigeons will detect a decrease in the reinforcement criterion. To illustrate, assume that after a large number of sessions N changes from 32 to 4. Initially, the run lengths will average approximately 32 pecks but, because $\delta > \alpha$, the effect of reinforcement when $X < \mu$ will be greater than the effect of reinforcement when $X \geq \mu$;

the net effect will be a decrease in μ[5]. However, both model and pigeon will have difficulties adjusting to large increases in the reinforcement criterion. If after a large number of sessions N changes from 4 to 32, neither the model nor a naive pigeon will generate run lengths sufficiently large to be reinforced; both will require successive approximations. How stimulus control assumptions may be added to the model in order to account for the data from Experiment 2 and from other studies (e.g., Platt & Johnson, 1971) remains to be determined.

REFERENCES

Alsop, B., & Honig, W. K. (1991). Sequential stimuli and relative numerosity discriminations in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 17*, 386–395.

Catania, A. C., & Reynolds, G. S. (1968). A quantitative analysis of responding maintained by interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior, 11*, 327–383.

Emmerton, J. (2001). Birds' judgments of number and quantity. In R. G. Cook (Ed.), *Avian visual cognition.* Retrieved March 17, 2007, from http://www.pigeon.psy.tufts.edu/avc/emmerton/

Fetterman, J. G. (1993). Numerosity discrimination: Both time and number matter. *Journal of Experimental Psychology: Animal Behavior Processes, 19*, 149–164.

Fetterman, J. G., Dreyfus, L. R., & Stubbs, D. A. (1985). Scaling of response-based events. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 388–404.

Galbicka, G., Fowler, K. P., & Ritch, Z. J. (1991). Control over response number by a targeted percentile schedule: Reinforcement loss and the acute effects of d-amphetamine. *Journal of the Experimental Analysis of Behavior, 56*, 205–215.

Gallistel, C. R. (1990). *The organization of learning.* Cambridge, MA: Bradford Books/MIT Press.

Hobson, S. L., & Newman, F. (1981). Fixed-ratio counting schedules: Response and time measures considered. In M. L. Commons, & J. A. Nevin (Eds.), *Quantitative analysis of behavior: Vol. 1. Discriminative properties of reinforcement schedules* (pp. 193–224). Cambridge, MA: Ballinger.

Keen, R., & Machado, A. (1999). How pigeons discriminate the relative frequency of events. *Journal of the Experimental Analysis of Behavior, 72*, 151–175.

Laties, V. G. (1972). The modification of drug effects on behavior by external discriminative stimuli. *Journal of Pharmacology and Experimental Therapeutics, 183*, 1–13.

Leak, T. M., & Gibbon, J. (1995). Simultaneous timing of multiple intervals: Implications of the scalar property. *Journal of Experimental Psychology: Animal Behaviour Processes, 21*, 3–19.

Machado, A., & Keen, R. (2002). Relative numerosity discrimination in the pigeon: Further tests of the linear-exponential-ratio model. *Behavioural Processes, 57*, 131–148.

Mechner, F. (1958). Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behavior, 1*, 109–122.

Mechner, F., & Guevrekian, L. (1962). Effects of deprivation upon counting and timing in rats. *Journal of the Experimental Analysis of Behavior, 5*, 463–466.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9*, 320–334.

Meck, W. H., Church, R. M., & Gibbon, J. (1985). Temporal integration in duration and number discrimination. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 591–597.

Platt, J. R., & Johnson, D. M. (1971). Localization of position within a homogeneous behavior chain: Effects of error contingencies. *Learning & Motivation, 2*, 386–414.

Rilling, M. (1965, April 23). Signal detection in fixed-ratio schedules. *Science, 148*, 526–527.

Rilling, M. (1967). Number of responses as a stimulus in fixed interval and fixed ratio schedules. *Journal of Comparative and Physiological Psychology, 63*, 60–65.

Rilling, M. (1993). Invisible counting animals: A history of contributions from comparative psychology, ethology, and learning theory. In S. T. Boysen, & E. J. Capaldi (Eds.), *The development of numerical competence: Animal and human models* (pp. 3–37). Hillsdale, NJ: Erlbaum.

Roberts, W. A. (1995). Simultaneous numerical and temporal processing in the pigeon. *Current Directions in Psychological Science, 4*, 47–51.

Roberts, W. A. (1998). *Principles of animal cognition.* New York: McGraw-Hill.

Roberts, W. A., & Mitchell, S. (1994). Can a pigeon simultaneously process temporal and numerical information? *Journal of Experimental Psychology: Animal Behavior Processes, 20*, 66–78.

Roberts, W. A., Roberts, S., & Kit, K. A. (2002). Pigeons presented with sequences of light flashes use behavior to count but not to time. *Journal of Experimental Psychology: Animal Behavior Processes, 28*, 137–150.

Whitaker, J. S., Lowe, C. F., & Wearden, J. H. (2003). Multiple-interval timing in rats: Performance on two-valued mixed fixed-interval schedules. *Journal of Experimental Psychology: Animal Behavior Processes, 29*, 277–291.

---

[5] The mean of the distribution will continue to decrease and approach *N* until the following stability condition is attained, $\frac{1}{2}(\alpha - \beta \times \mu) = \delta \times \int_N^\mu f(\tau, \mu, \gamma\mu)d\tau$, where f($\tau$,$\mu$,$\gamma\mu$) is the normal density function with mean μ and standard deviation $\gamma\mu$ evaluated at τ.

To obtain an equation for the curves on the right panels of Figures 5, 6, 7, and 8, we note that each corresponds to the probability distribution of run length on the left key at the moment the pigeon switches to the right key. Let $P(R = n)$ represent the probability that run length equals n. To derive an expression for $P(R = n)$, we note that $R = n$ means that one of two events happened, either $r_1 = n$ or $r_1 + r_2 = n$, because the peck on the right key ended either a first run, in which case $r_1 = n$, or a second run, in which case $r_1 + r_2 = n$. Hence,

$$P(R = n) = P(''r_1 = n'' \text{ Or } ''r_1 + r_2'' = n)$$
$$= P(r_1 = n) + P(r_1 + r_2 = n), \quad (A1)$$

where the last equality follows because the events ''$r_1 = n$'' and ''$r_1 + r_2 = n$'' are mutually exclusive.

The first term, $P(r_1 = n)$, may be approximated by a Gaussian density function with mean μ_1 and standard deviation σ_1:

$$P(r_1 = n) \approx f(n, \mu_1, \sigma_1). \quad (A2)$$

The second term, $P(r_1 + r_2 = n)$ can be obtained by conditioning on the length of r_1. That is,

$$P(r_1 + r_2 = n)$$
$$= \sum_{i=0}^{\min(n-1,15)} P(r_1 + r_2 = n | r_1 = i) P(r_1 = i) \quad (A3)$$
$$= \sum_{i=0}^{\min(n-1,15)} P(r_1 = i) P(r_2 = n - i | r_1 = i).$$

Note that the summation index did not extend beyond 15 because all first runs of length greater than 15 were reinforced and ended the trial; there were no r_2 runs under those circumstances. In contrast, for the sum of the two runs to equal n, the first run could not exceed n – 1 pecks (the second run could not equal 0).

In the preceding summation, one must consider the special case $i = 4$ and distinguish two events, namely, the first run was reinforced (which occurred with probability p), or it was not (probability $1 - p$). In the first event, the trial ended and there was no opportunity for a second run; in the second event, the trial continued and r_2 was produced. Hence, when i

$= 4$ only nonreinforced trials contribute to the summation:

$$P(r_1 + r_2 = n)$$
$$= \sum_{i=0, i \neq 4}^{\min(n-1,15)} \{P(r_1 = i) P(r_2 = n - i | r_1 = i)\} \quad (A4)$$
$$+ P(r_1 = 4) P(r_2 = n - 4 | r_1 = 4)(1 - p).$$

After multiplying and rearranging one gets:

$$P(r_1 + r_2 = n)$$
$$= \sum_{i=0}^{\min(n-1,15)} \{P(r_1 = i) P(r_2 = n - i | r_1 = i)\} \quad (A5)$$
$$- P(r_1 = 4) P(r_2 = n - 4 | r_1 = 4) p.$$

If one assumes independence between r_1 and r_2, then Equation A5 simplifies to:

$$P(r_1 + r_2 = n)$$
$$= \sum_{i=0}^{\min(n-1,15)} \{P(r_1 = i) P(r_2 = n - i)\} \quad (A6)$$
$$- P(r_1 = 4) P(r_2 = n - 4 | r_1 = 4) p.$$

In Equation A6, the term in brackets represents the convolution of r_1 and r_2, and the last term represents the correction for the case $r_1 = 4$. We approximate the distribution of r_2 by a Gaussian density function, as we did for r_1:

$$P(r_1 + r_2 = n)$$
$$= \sum_{i=0}^{\min(n-1,15)} \{f(i, \mu_1, \sigma_1) f(n - i, \mu_2, \sigma_2)\} \quad (A7)$$
$$- f(4, \mu_1, \sigma_1) f(n - 4, \mu_2, \sigma_2) p.$$

Finally, combining Equations A2 and A7 yields

$$P(R = n) = f(n, \mu_1, \sigma_1)$$
$$+ \sum_{i=0}^{\min(n-1,15)} \{f(i, \mu_1, \sigma_1) f(n - i, \mu_2, \sigma_2)\} \quad (A8)$$
$$- f(4, \mu_1, \sigma_1) f(n - 4, \mu_2, \sigma_2) p.$$

Equation A8 was used to fit the data in the right panels of Figures 5, 6, 7, and 8.